2006

Designing Computer Agents with Facial Personality to Improve Human-Machine Collaboration

Brian Esley Tidball
Wright State University

DESIGNING COMPUTER AGENTS WITH FACIAL PERSONALITY TO IMPROVE HUMAN-MACHINE COLLABORATION

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science in Engineering

By

BRIAN ESLEY TIDBALL
B.S.E. University of Washington, 1998

2006
Wright State University

WRIGHT STATE UNIVERSITY

SCHOOL OF GRADUATE STUDIES

March 17, 2006

I HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER MY SUPERVISION BY Brian Esley Tidball ENTITLED Designing Computer Agents with Facial Personality to Improve Human-Machine Collaboration BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF Master of Science in Engineering.

_____
Jennie J. Gallimore, Ph.D.
Thesis Director

_____
S. Narayanan, Ph.D.
Department Chair

Committee on
Final Examination

_____
Jennie J. Gallimore, Ph.D.

_____
Raymond R. Hill, Ph.D.

_____
Daniel Voss, Ph.D.

_____
Joseph F. Thomas, Jr., Ph.D.
Dean, School of Graduate Studies

ABSTRACT

Tidball, Brian Esley M.S.E., Department of Biomedical, Industrial and Human Factors Engineering, Wright State University, 2006. Designing Computer Agents with Facial Personality to Improve Human-Machine Collaboration.

The development of computer agents to enhance human-computer interfaces is an evolving field of study. This study examined whether people perceive personality in static digital faces that portray expressions of emotion, and if the digital faces would influence human performance on a simple human-machine collaborative task.

The first experiment measured user-perception of personality based on the emotional expression in two sets of five static digital faces, one face from each set represented the five primary emotions (Anger, Fear, Joy, Sad and Neutral). The independent variables were: emotional Expression, Personality Trait, Face Set, Gender and Nationality. Using an internet survey, subjects were asked to rate the faces on the 30 personality sub-traits of the Big-Five Factor personality model on a scale of 1 to 5 (1 = not characteristic, 3 = partially characteristic, 5 = fully characteristic). These ratings were used to compare the perceived personality of each face. The results from this first phase revealed that participants provided different ratings of the personality sub-traits based on the emotional expression of a static digital face indicating perception of personality based on expression. There were also some isolated differences related to Gender and Face Set.

Based on the results from experiment one, three faces were chosen for experiment two. The second experiment measured how faces with identified personality traits influence decision making in a simple collaborative task. Subjects were asked to read a survival scenario and then rank a list of items according to their value for survival.

During the task, a computer agent represented by a static digital face would make two suggestions in text form. The independent variables were Face, Suggestion Accuracy (Good, Moderate, Poor), Gender and Nationality (American, Indian). The dependent variables were Task Performance Score and number of Suggestions Used. The results revealed that the different Faces did not have a significant impact on either Task Score or the Suggestions Use. American subjects had better Task Scores than Indian subjects. This might be due to the Indian subject's lack of familiarity with the survival scenario task. Female participants were more likely to use the suggestions provided by the computer agent. As expected, the accuracy of the suggestions influenced both task performance and suggestion use. When the Suggestion Accuracy was Good, task score improved and the number of suggestions used increased. There were no differences when Suggestion Accuracy was Moderate or Poor, indicating subjects were taking care in performing the task.

The lack of significant differences based on the use of a static digital face to represent a computer agent was likely caused by the lack of required interaction with the computer agent. While the agent was pictured on the screen the task did not require the user to interact with it or look at the suggestion. Future research considerations are discussed.

TABLE OF CONTENTS

Page

TABLE OF CONTENTS continued

LIST OF FIGURES

Figure Page

LIST OF FIGURES continued

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to acknowledge the guidance, support and insight provided by Dr. Jennie J. Gallimore, Dr. Raymond R. Hill and Dr. Daniel Voss, without whom this project would not have been possible. Additionally, I would like to thank all the individuals who volunteered their time to participate as research subjects, allowing me to collect the data necessary to complete this project. And most importantly, I would like to thank my wife for her love, support and encouragement throughout this process.

1. INTRODUCTION

The growing size and complexity of computer systems is spawning the use of innovative ways to better collect, filter, analyze and present information to reduce the cognitive demands on human operators. One such innovation is agent-based software. Software agents are designed as "smart" programs that serve the needs of a human operator by filtering or searching out information for the operator. Agents are designed and programmed for these specific tasks to simplify the interaction between the human and the computer by parsing responsibilities and protecting the human from vast amounts of irrelevant data. As computer systems continue to grow, it becomes increasingly important to provide more natural human-computer interactions (HCI) to reduce cognitive load. One possible improvement is to design agents with personality, thus making the interaction feel like a collaborative partnership that augments human capabilities instead of simply providing an interface to display information. The development of software agents with personalities is potentially the next evolutionary step in human-computer interfaces.

This thesis will focus on the human perception of personality, based on digital facial representation. It will then measure the effect of the personality on influencing human decision-making in human-machine collaborative environments.

Because there is limited research on both software agents with personality and classifying personality TRAITS through facial expression it is critical to introduce several topics related to this research before proceeding. These topics described below include: collaboration, personality theory, physiognomy, emotion, facial EXPRESSIONS, augmented cognition and computer agents with personality.

2. BACKGROUND

2.1 Collaboration

Collaboration is the process of "jointly working with others…especially in an intellectual endeavor" (*Merriam-Webster Online*). Although researchers focus attention on different aspects of collaboration, there is agreement that as task demands exceed the abilities of an individual collaboration between agents becomes critical (Baker and Salas, 1992). Research on collaboration in complex systems highlights ten key characteristics for success (Baker and Salas, 1992; Cannon-Bowers and Salas, 1998; Hoc, 2001):

- Effective communication
- Shared mental model of the system
- Converge on decisions reached by consensus
- Evaluate and adapt to changes
- Take actions to solve problems
- Anticipate others actions and needs
- No power struggle
- Clear objectives
- Assignments are clear and accepted
- Exchange of information

Good communication within a team that results in the accomplishment of necessary tasks is elemental to this list of characteristics as well as any study of collaboration.

Generally collaboration occurs between humans; but, according to the Computers as Social Actors (CASA) paradigm, the social rules guiding human-human interactions can be applied to human-machine interactions (Nass, Steuer, Tauber and Reeder, 1993). Hoc (2001) inherently supports this idea when he states that human-machine collaboration can be enhanced through improvements in user interface design, the use of

expert systems and by addressing the cognitive and social aspects of human-human collaboration.

Decision-making roles are another key aspect of successful collaboration and can be addressed vertically (hierarchical) or horizontally (heterarchical). An air traffic control study showed that horizontal decision making was best in a situation where each actor makes decisions independently and collaborates only when situations arise (Millot and Lemoine, 1998). Other situations may call for vertical decision-making where the computer agent is only responsible for making suggestions, placing the human in the position of authority. As designers begin to build machines or agents as collaborative partners, researchers are starting to address the key issues of decision-making roles, specifically credibility and trust. Fogg and Tseng (1999) stated that credibility is essential for effective collaboration and divided credibility into four categories: device, interface, functional and information credibility. Humans view machines as more credible when they are predictable and dependable (Muir, 1987). This view can be enhanced with labels appropriate to their expertise similar to the titles we give humans, such as Doctor or Professor (Reeves and Nass, 1996). It is also important to note that Nass, Fogg and Moon (1996) observed that Similarity-Attraction Theory (attraction to like personalities) improved machine credibility.

Although many of the studies in human-machine collaboration have a narrow scope, many of the rules that guide human-human collaboration apply to human-machine interaction. Research is still necessary to determine if computer agents with personality improve collaboration and to find ways of accurately incorporating personality into these computer agents.

2.2 Personality Theory

Personality is the set of "characteristics that distinguishes…the totality of an individual's behavior and emotional characteristics" (*Merriam-Webster Online).* The vast number and complexity of characteristics available to describe a personality seem limitless. To help manage the complexity, personality theories or models employ methods to organize and simplify these characteristics in order to classify the personality of individuals. Winter and Barenbaum (1990) make the distinction between four types of personality models: motivational, cognition, social context and trait. Motivational models seek to identify behavior patterns aimed at a goal, such as Maslow's Hierarchy of Needs (1970). Cognition models seek to understand the constructs of how the individual perceives and interacts with the environment. Social context models examine individual traits as they relate to interactions in a group. Trait models seek to identify a small list of factors that in combination account for the majority of the variation in an individual's personality. A common example of a trait model is Goldberg's Big-Five Factor model (1999).

Sir Francis Galton is credited with recognizing the fundamental lexical hypothesis which states that you can identify "the more conspicuous aspects of the character by counting in an appropriate dictionary." Simply stated, personality traits are a part of our normal vocabulary, with more words available to describe more important traits. Galton (1884) also surmised that although there are a thousand subtly unique words used to describe character, each word shares a large part of it's meaning with many others. L.L. Thurstone (Goldberg, 1990) was one of the first to develop trait-based factor analysis based on Galton's ideas. Thurston's research found that "five factors are sufficient to

account for the coefficients" (1934) or variability of 60 personality traits. Since then, a number of researchers (Borgatta (1964), Digman and Takemoto-Chock (1981) and McCrae and Costa (1985)) have analyzed sets of more than a thousand words and reported similar five-factor structures, where every word could be grouped within one of five categories or "factors".

Goldberg's Big-Five Factor model is a trait model based on this five-factor structure. Each of the five TRAITS (I. Extroversion, II. Agreeableness, III. Contentiousness, IV. Emotional Stability vs. Neuroticism, and V. Intellect or Openness) are divided into six sub-traits. (See Appendix B, for a complete list of sub-traits and definitions.) Considerable research has validated several aspects of the Big-Five Factor Model. Solds and Vaillant (1999) validated the stability of the five factors over time and across domains. McCrae and Costa (1997) validated the five factors across six different cultures. These factors also represent individual differences in approaches to problem solving (Buss, 1996; McCrae and Costa, 1999). Other efforts to distinguish additional factors have resulted in the realization that the traits are already represented by the Big-Five Factor model.

Although the Big-Five Factor model has been widely accepted and used to classify and discuss personality, it has only recently been applied to the development of computer agents with personality (Prabhala and Gallimore, 2005). This study will use the 30 sub-traits from the Big-Five Factor model to measure perceived traits in the facial representations of computer agents.

2.3 Physiognomy

Physiognomy is the art of reading personality traits from the characteristics of static faces. This course of study dates back to ancient Greece and the writings of Aristotle (Zebrowitz, 1997), though it received more significant consideration with the publication of *Essays on Physiognomy* by Johan Caspar Lavater in the late 18th century. It continues to be a popular area of study today.

In a review of physiognomy research (Hassin and Thrope, 2000) we see that there are three primary findings. First, evidence suggests that people can and do infer personality traits from faces. Second, different people reliably infer the same personality traits from given faces, including cross culturally. Third, the validity of these judgments is questionable. Early research shows no significant correlation between physiognomic inferences and actual personality traits. More recent studies focusing on general impression rather than discrete traits show that there may be a correlation, though the methods of these studies may be confused with behavioral aspects of the individuals.

Recent research by Hassin and Thrope (2000) hypothesizes that "physiognomy is an integral part of social cognition." They refer to this as the effect of "reading from and into faces." The conclusions they draw from a series of six studies include: physiognomic information changes people's impression of information, where more ambiguous information is interpreted with more reliance on facial impression. Physiognomic information is incorporated in decision making, even when asked to intentionally ignore people's faces. Furthermore, individual confidence in their physiognomic judgments far exceeds the accuracy of those interpretations.

Physiognomy may be an important consideration in the design and use of facial representation in computer agents. Research is needed to determine if these interpretations are similarly applicable to computer agents. If transferable, careful selection of facial features may be used to improve the effectiveness of human interaction with these agents by making them more genuine. The validity of these judgments becomes a design consideration and a measure of intended personality.

2.4 Emotion and Facial Expressions

The study of how people project and perceive emotion through facial EXPRESSIONS has been and continues to be intensely examined. Until recently this research has been confined by the use of exaggerated static expressions with the neglect of facial features and situational cues (Hagar and Ekman, 1983). Despite these limitations, the research shows that humans are universally and cross culturally proficient at expressing and interpreting five primary emotions: anger, fear, joy, sadness and disgust (Ekman, Sorenson and Frieson, 1969).

Emotions such as surprise and shame can be proficiently expressed and interpreted, but Izard (1971) concludes that head position is more revealing than facial expression. Inconclusive results on evaluating other emotions are attributed to an individual's inability to accurately project complex emotions and further confounded by our inability to reliably interpret the intended emotion (Ekman, 1979). Additionally, individual interpretation of expressions is dependent on the mood of the observer (Ruckmick, 1921) and the intensity of the expression. Recent biological and neurological studies of emotion have shown that there is a relationship between facial expression and

autonomic measures of arousal (Hagar and Ekman, 1983) supporting earlier ideas that many facial EXPRESSIONS are innate and not visually learned (Charlesworth and Kreutzer, 1973). This supports conclusions on the universal nature of the five primary facial EXPRESSIONS.

Much of the research on emotion has relied on the use of static pictures similar to those developed by Ekman for his work in the late 1960s. These photographs were produced by actors expressing a designated emotion, though a small number of researchers, including Ekman in later works, have used elicited expressions as well as more elaborate methods to track and study dynamic changes. A growing number of researchers believe many additional emotions may be accurately conveyed and interpreted through the dynamic change in expression (duration, recurrence, intensity or combined expressions) or the use of contextual clues (setting, posture or dialog). Unfortunately, results are generally inconclusive due to the complexity of these actions and interactions (Hagar, 1983).

To help deal with the complexity of facial expression, Ekman and Freiesen (1976) developed the Facial Action Coding System (FACS) to enable the measurement of all visible facial movements. Although FACS is limited by the observer's ability to discriminate the movements of 46 "action units," it allows the description of all expressions using muscle actions; therefore it is not reliant on emotional labels and helps overcome problems due to physiognomic differences. Using this system we can recognize that there are hundreds of thousands of visibly distinguishable static facial EXPRESSIONS, though Ekman admits that many facial actions are not related to an emotion. Using this method to score pictures of emotion, there appear to be hundreds of

muscle combinations that convey emotional meanings. Research is still needed to determine if naïve observers could distinguish between the various connotations of particular emotions. Expanding on this system, Essa and Pentland (1997) and others have developed computer vision systems capable of dynamically tracking, coding and analyzing facial movements with detailed precision. These methods move beyond static analysis and remove the need for experts to tediously identify each muscle movement. However, research is still needed to look at interpretation and recognition of these more detailed and dynamically tracked EXPRESSIONS.

Research shows that people can and do read both emotion and personality from other peoples' faces. Recognizing that people reliably identify the five primary emotions, it may be possible to design faces for software agents that universally project emotions which translate to desirable personality TRAITS for collaborative partners.

2.5 Augmented Cognition

The emerging field of Augmented Cognition (AugCog) seeks to develop an interactive human-computer system where the state of the human is automatically measured, analyzed and adapted to, to improve cognitive performance (Schmorrow and Kruse, 2002). These systems are comprised of software and hardware tools that work to bridge the gap between the human brain and the system to improve the task handling capabilities of the operator by sensing cognitive thresholds (Eitelman, Wheeler-Atkinson, Walwanis-Nelson and Stiso, 2005). There are numerous biological and behavioral measures under investigation: attention, electroencephalograph (EGG), error rate, eye tracking, heart rate, memory, near infrared spectroscopy (NIRS), stress and workload.

While some research is trying to determine which measures are effective, other groups like Johnson, Kulkarni, Raj, Carff and Bradshaw (2005) are working to combine multiple sensors in an Adaptive Multi-Agent Integration (AMI) framework to link these measures and provide a comprehensive understanding of the state of the human operator.

Once the state of the operator is determined, an AugCog system automatically adapts itself to user needs by providing mitigation and mediation of the interface to optimize cognitive performance (Schmorrow and Kruse, 2002). As cognitive demands increase, AugCog systems could utilize graphics, text or audio to direct user attention to new, changing or priority information. This could be accomplished by hiding or fading less critical information, using data summary, or highlighting (bigger, brighter or flashing) important information. If the operator is approaching overload when decision points occur, the system could recommend alternatives and include future impacts, reducing the operator's need to derive them, thereby reducing cognitive load.

In the future Augmented Cognition could be adapted to control the personalities of computer agents to help mitigate cognitive load. Agent personality or emotion could be altered to gain attention (raise voice, get excited, encourage, reprimand, etc.) or inform the user that they (the agents) will handle tasks that are distracting or causing overload.

3. RELATED RESEARCH

3.1 Computer Agents with Personality

Complex systems that require human operators to work with automated agents are a form of human-computer collaboration. The growing support for improving this collaboration by drawing on the strengths of human-human collaboration, begs the question: How do we make these agents more human? One answer may be to give the agents personality. Recent work by Prabhala and Gallimore (2005a, 2005b and 2005c) explores the actions, language and behaviors that signify personality TRAITS (from the Big-Five Factor model) that are important in collaboration. For this study, they blocked participants into three groups: computer team member, existing team member and ideal team member. Each participant began by completing a Big-Five Factor personality test and then, based on their group, was asked to rate and describe a team member or members in relation to the 30 sub-TRAITS in the Big-Five Factor model. Results from all three groups showed significant trends in the ratings of both central and sub-TRAITS in the Big-Five Factor model. There was high correlation between an ideal team member and existing team members. Ideal team members are rated and described as extraverts, agreeable, contentious and not neurotic and within the central trait of openness. Intellect and imagination were identified as the only two desirable sub-TRAITS. Results specifically from the computer team member block, provided a number of useful insights:

- Subjects perceived personality in computer team members
- Personalities were perceived to be different
- Different subjects had different impressions
- Subjects could identify the actions, language and behaviors that led to their impressions
- There was no significant difference across culture or gender

In addition to these conclusions, Prabhala and Gallimore were able to capture the actions, languages and behaviors associated with perceptions of personality. This data is critically important for the further development of computer agents with personality. By categorizing these items, it will be possible to incorporate them into human behavioral models to make human-computer collaboration more realistic.

3.2 Providing a More Natural HCI

Many researchers (Pew and Mavor 1998, Silverman, Cornwell and O'Brien 2003, and Wray and Laird 2003) have acknowledged the need to develop agents that act more like the human operators, to provide a more natural HCI. The hope is that a collaborative relationship will enhance trust in the system and reduce cognitive workload. Unfortunately, there is limited research on computer agents as they relate to personality, emotion, facial representation and facial expression. The question remains: how to effectively incorporate personality into agents and determine whether personality can be conveyed through facial EXPRESSIONS. Trappl and Petta's 1997 book, *Creating Personalities for Synthetic Actors,* described an agent's personality through the use of visual and verbal stereotypes. Research by Prabhala and Gallimore (2005) looked at whether participants could perceive personality characteristics in a computer game. Results showed that different personalities were perceived, but more importantly they documented the actions, language and behaviors that led to their perceptions.

4. OBJECTIVES

The principal foci of this study was to measure whether subjects perceive personality in static digital faces and determine if static digital faces presented as partners in a collaboration task will influence performance of the collaboration task. To explore these ideas, two experiments were conducted. The first experiment was designed to determine if participants perceive personality TRAITS based on the emotional expression of a static digital face (Phase I). The second experiment was designed to determine if the faces with identified personality TRAITS influence decision making in a collaborative task (Phase II).

5. PHASE I: HYPOTHESES

The primary hypothesis for the first phase of this study is concerned with participant's perception of personality TRAITS in the digital facial representations. Within this primary hypothesis six specific hypotheses will be tested:

Table 1: Phase I - Hypothesis and Expectations

	Null Hypotheses	Expectation
1.	There will be no difference in the ratings based on **Nationality** (main effect of Nationality).	Fail to Reject Hypothesis
2.	There will be no difference in ratings based on **Gender** (main effect of Gender).	Reject Hypothesis
3.	There will be no difference in ratings based on **FaceSet** (interactions with FaceSet).	Reject Hypothesis
4.	There will be no difference in ratings based on emotional **Expression** (main effect of Expression).	Reject Hypothesis
5.	There will be no difference in ratings among **individual faces** (interaction of FaceSet and Expression).	Reject Hypothesis
6.	There will be no difference in ratings of the five **Personality Traits** (Extroversion, Agreeableness, Contentiousness, Neuroticism and Openness) based on expression (interaction of Expression and Personality Traits).	Reject Hypothesis

Testing the hypotheses listed in Table 1 will demonstrate if there are significant differences or similarities in the way people perceive each digital face and emotional expression. By grouping and averaging the subjects ratings related to the independent sub-traits into their central personality trait groups as described by the Big-Five Factor personality model, a mean score for each of the five central personality traits was determined for all 10 faces. These scores were used to conduct the analysis necessary to test these hypotheses.

6. PHASE I: METHOD

6.1 Experimental Design

The experiment for the first phase was designed to present a series of stimuli each of which participants rated each stimuli on 34 characteristic terms. A participant completed the questionnaire only once, representing one experimental trial. Within each questionnaire or trial, subjects were presented with one set of five facial stimuli. The experimental design for the first phase of this study was a 30 x 5 x 2 x 2 x 2 mixed factorial design. The within-subject variables are 30 personality sub-trait characteristics (PERSONALITY TRAITS) defined in Appendix B and facial EXPRESSIONS (Anger, Fear, Joy, Sad, Neutral). The between-subject variables are FACE SETS (A, B), participant NATIONALITIES (American, Indian) and participant GENDER (Male, Female).

6.2 Subjects

Subjects were solicited via an email advertisement sent to all engineering students at Wright State University. Fifty-eight participants completed the entire survey. Demographic data showed that nationality was divided into 48 American participants and 10 Indian participants, with gender more evenly distributed between 30 males and 28 females. Nearly all participants were under the age of 30. The web survey randomly blocked participants into two groups: Face Set A (28) and Face Set B (30).

6.3 Stimuli

Two sets of faces (A and B) consisting of five faces each (10 faces total) were used as stimuli to measure human perception of digital faces in relation to PERSONALITY TRAITS. Each set of faces was static and comprised of five emotional EXPRESSIONS: anger, fear, joy, sadness and neutral. These EXPRESSIONS were selected to align with the emotions from Ekman's (1969, 1979, 1983) well-supported research showing that these emotions are universally and proficiently recognizable. The expression of disgust was excluded because it was not available for set A and looked identical to anger when created for set B.

The first set of faces (A) was borrowed from the research of Goren and Wilson (2006). They generated the faces using an averaging and filtering procedure described by Wilson, Loffler and Wilkenson (2002). The faces were produced from a database of facial measurements from 37 individuals, averaged into a single face and then bandpass filtered. The result is the neutral face shown in Figure 1.

Figure 1: Face A5-neutral expression

This neutral face was then digitally manipulated into the four EXPRESSIONS, shown in Figure 2.

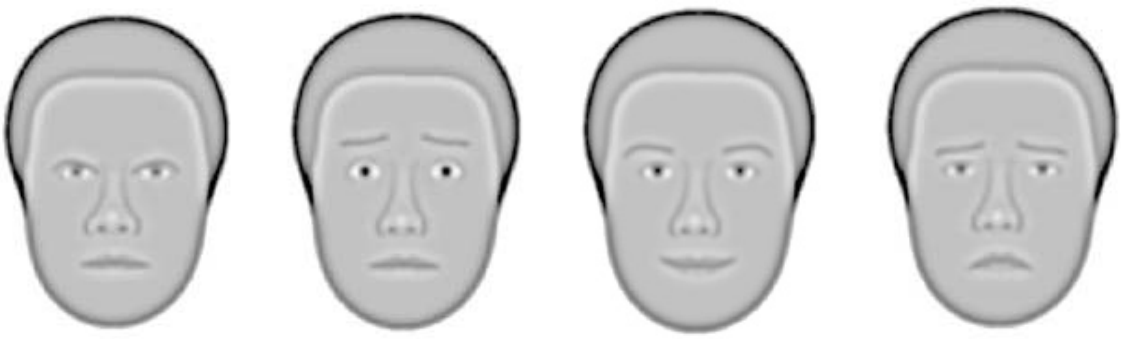


Figure 2: Faces, A1-anger, A2-fear, A3-joy and A4-sadness

The second set of faces (B) was drawn using FaceGen Modeller 3.1, software package. The faces were generated by making minor adjustments to the settings for an "average face" of a 30-year-old male with 50% European "race morphing." The result is the neutral face shown in Figure 3.

Figure 3: Face B5 - neutral expression

This face was then adjusted using the "morph" and "shape" tools in the FaceGen software to produce the desired EXPRESSIONS. The following settings were used to produce the EXPRESSIONS shown in Figure 4:

- Anger – "Anger" (0.5)
- Fear – "Fear" (0.5) and "Surprise" (0.5)
- Joy – "Smile Closed" (0.5)
- Sadness – "Sad" (0.75) and "Symmetric/Mouth - Happy/Sad" (5.0)

Figure 4: Faces, B1-anger, B2-fear, B3-joy and B4-sadness

6.4 Apparatus

The FaceGen Modeller 3.1 software used to develop the second set of stimuli (face set B) was run on a 800MHz desktop personal computer with 512Mb memory. Files were saved as a "*.jpg" picture file for use in the online survey and visual basic.NET program written for Phase II.

Qualtrics.com online survey software provided survey development tools, hosting and data collection. The Qualtrics.com development tools provided the functionality necessary to randomize the order of the stimuli and randomize the question order

between stimuli. Using a web survey, participant hardware could not be controlled except to require the use of a computer with internet access and a web browser.

6.5 Procedure

The first phase of research collected data through the distribution of an online questionnaire (See Appendix C) using Qualtrics.com online survey software. This questionnaire opened with a consent form followed by demographic questions to document the nationality, age, gender and education background of the participants. People who rejected the consent document were denied access to the questionnaire. The main body of the survey consisted of a series of similar pages showing one of the ten static faces described above, followed by a list of characteristics. The characteristics are the collection of the 30 personality sub-TRAITS from the Big-Five Factor personality model (Goldberg, 1990) and the five primary emotions described by Ekman's research. Each participant was randomly blocked into groups A and B, and was only presented with the five faces from the corresponding face set (A or B). The participant was then asked to rate each of the five faces on all 34 characteristics using a five-point Likert scale (1 = Not Characteristic, 3 = Partially Characteristic, 5 = Fully Characteristic) as shown in Appendix C. The 34 characteristics were randomly divided into three groups to limit the number of characteristics per screen; this ensured that the face and all terms were simultaneously visible without scrolling. To minimize order effects, the order of the faces, the order of the three groups of terms and the order of the terms on each page were each randomized. The Qualtrics.com survey website automatically collected and saved the responses for each participant.

6.6 Dependent Variable

Each subject was presented with five faces and asked to rate each face on all 34 characteristic terms, totaling 170 responses per subject. These ratings represent the dependent variable for Phase I. Each participant represented a single replication of between-subject variables.

7. PHASE I: RESULTS

Subject ratings were analyzed using the JMP IN 5.1 statistical software package. The 58 subjects, each rated five faces on 30 different personality sub-TRAITS, producing 8,700 individual data points for the RATING dependant variable. On a five-point Likert scale a RATING of 1 corresponds to "not characteristic," 3 corresponds to "partially characteristic" and 5 corresponds to "fully characteristic." To conduct the analysis of this data the ratings for the six sub-traits for each of the five central PERSONALITY TRAITS identified in the Big-Five Factor personality model, were averaged to determine a mean RATING for each central PERSONALITY TRAIT. This is consistent with the way sub-trait scores are averaged into personality traits for the Big-Five Factor personality model.

Data analysis started with a full factorial ANOVA (GENDER * FACE SET * EXPRESSION * PERSONALITY TRAIT). The results of this ANOVA are shown in Table 2. Using a significance criterion of 0.05 we see that there are several significant interactions and main effects (Table 2). Simple-effects F-Tests were used to analyze significant interactions and the Tukey-Kramer Honestly Significant Difference test (Tukey HSD) was used to assess the significant main effects. The detailed analysis first looked at the complex interactions, then at the significant two-way interactions and finally looking at the significant main effects.

Table 2: ANOVA summary of subject's ratings when rating the face for personality trait

Between Subject	df	SS	MS	F Ratio	Prob > F
Gender	1	94.7821	94.7821	3.0529	0.0863
Face Set	1	127.0270	127.0270	4.0915	0.0481
Gender*Face Set	1	19.9913	19.9913	0.6439	0.4258
Subject[Gender,Face Set]	54	1676.5200	31.0467		
Within Subject					
Expression	4	560.9430	140.2360	40.9827	<.0001
Personality Trait	4	118.8620	29.7154	11.0220	<.0001
Gender*Expression	4	42.8696	10.7174	3.1321	0.0157
Gender*Personality Trait	4	8.9929	2.2482	0.8339	0.5049
Face Set*Expression	4	16.6897	4.1724	1.2194	0.3035
Face Set*Personality Trait	4	71.1351	17.7838	6.5964	<.0001
Expression*Personality Trait	16	1578.7800	98.6740	59.8299	<.0001
Gender*Face Set*Expression	4	23.1672	5.7918	1.6926	0.1528
Gender*Face Set*Personality Trait	4	3.6073	0.9018	0.3345	0.8545
Gender*Expression*Personality Trait	16	55.8779	3.4924	2.1176	0.0063
Face Set*Expression*Personality Trait	16	99.0677	6.1917	3.7543	<.0001
Gender*Face Set*Expression*Personality Trait	16	26.6399	1.6650	1.0096	0.4437
Subject*Expression[Gender,Face Set]	216	739.1140	3.4218		
Subject*Personality Traits[Gender,Face Set]	216	582.3350	2.6960		
Subject*Personality Traits*Expression[Gender,Face Set]	864	1424.9500	1.6492		

7.1 Three-way Interactions

FACE SET x EXPRESSION x PERSONALITY TRAIT

There was a significant three-way interaction among FACE SET, EXPRESSION and PERSONALITY TRAIT, $F(16, 864) = 3.7543$, $p \leq 0.0001$, as illustrated in Figure 5. A visual examination of this figure indicates different rating patterns for the traits across the different expressions. To further analyze the significance of this interaction a simple-effects F-Test by EXPRESSION was conducted. Results indicated a significant two-way interaction between FACE SET and PERSONALITY TRAIT for three of the five facial EXPRESSIONS: Anger ($F(4, 216) = 3.1940$, $p = 0.0142$), Sad ($F(4, 216) = 12.5217$, $p < 0.0001$) and Fear ($F(4, 216) = 3.0784$, $p = 0.0171$) as illustrated in Figures 6, 7 and 8. The two-way interaction was not significant for the Joy or Neutral EXPRESSIONS. When

EXPRESSION was Anger, the RATING for Extraversion was significantly greater for the FACE SET B (Figure 6). When the EXPRESSION was Fear the average RATING for each PERSONALITY TRAIT was greater for FACE SET B, with the exception of Neuroticism. When the EXPRESSION was Sad the average RATING for Extraversion and Contentiousness was higher for FACE SET B, while Neuroticism's RATING was greater for FACE SET A.

This simple-effects F-Test by EXPRESSION also showed that the effect of FACE SET was significant for the Neutral and Sad EXPRESSIONS: $F(1, 54) = 5.4391$, $p = 0.0234$ and $F(1, 54) = 5.3222$, $p = 0.0249$ respectively. As illustrated in Figure 9, when the EXPRESSION was Neutral or Sad the average RATING was significantly higher for FACE SET B. Although the other levels of EXPRESSION were not significant they appear to follow the same trend, reinforcing the idea that participants perceived the stimuli in FACE SET B to be more characteristic than for FACE SET A.

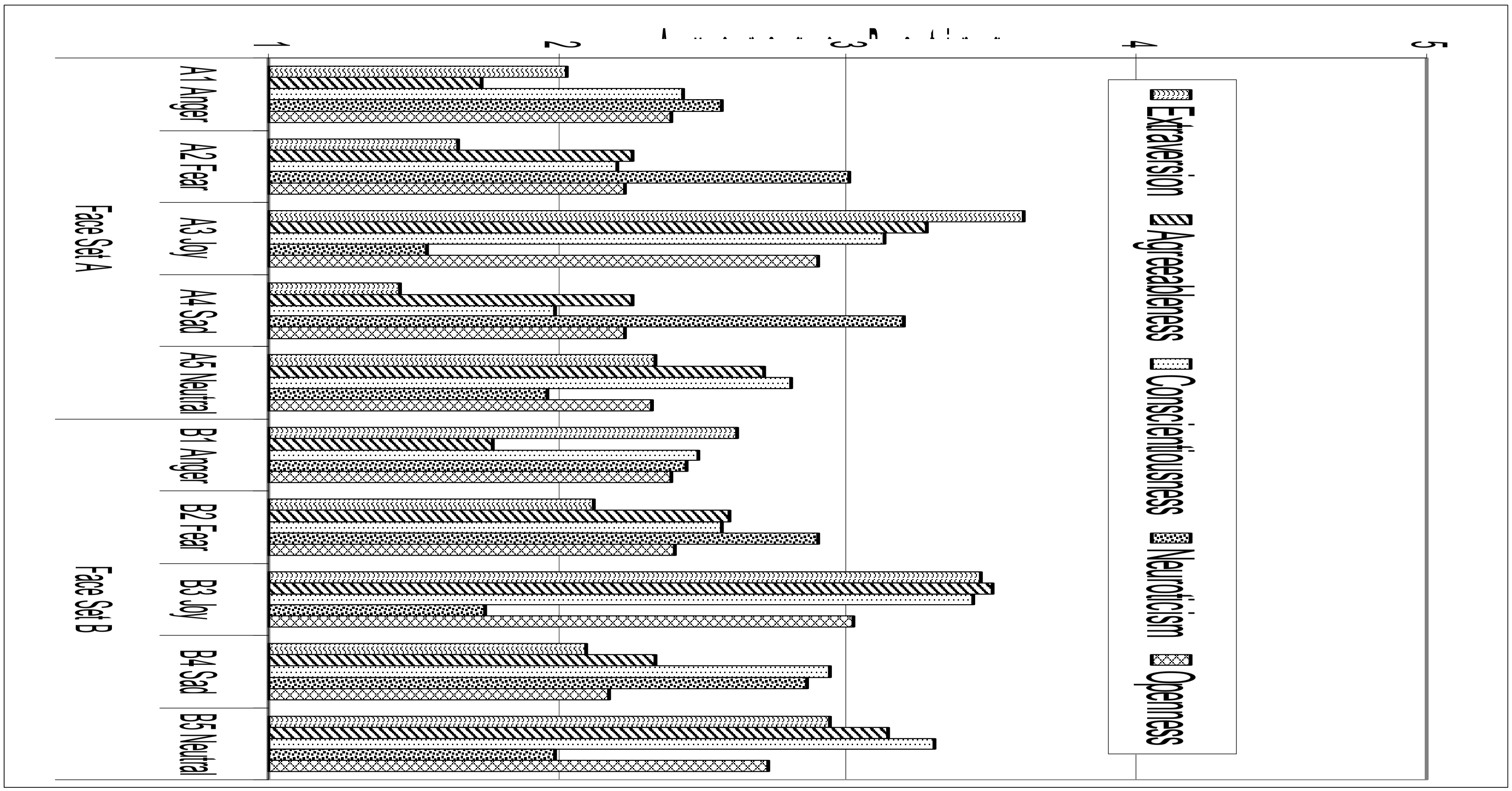


Figure 5: Average RATING for PERSONALITY TRAIT by FACE SET and EXPRESSION

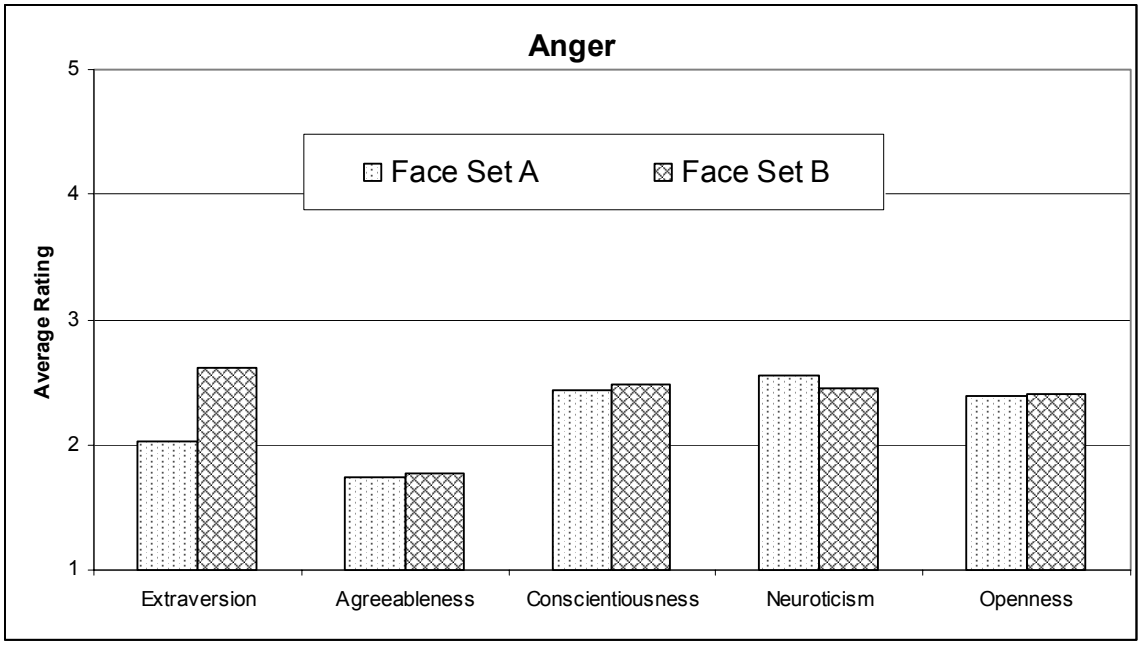


Figure 6: Average RATING for PERSONALITY TRAIT by FACE SET for the EXPRESSION Anger

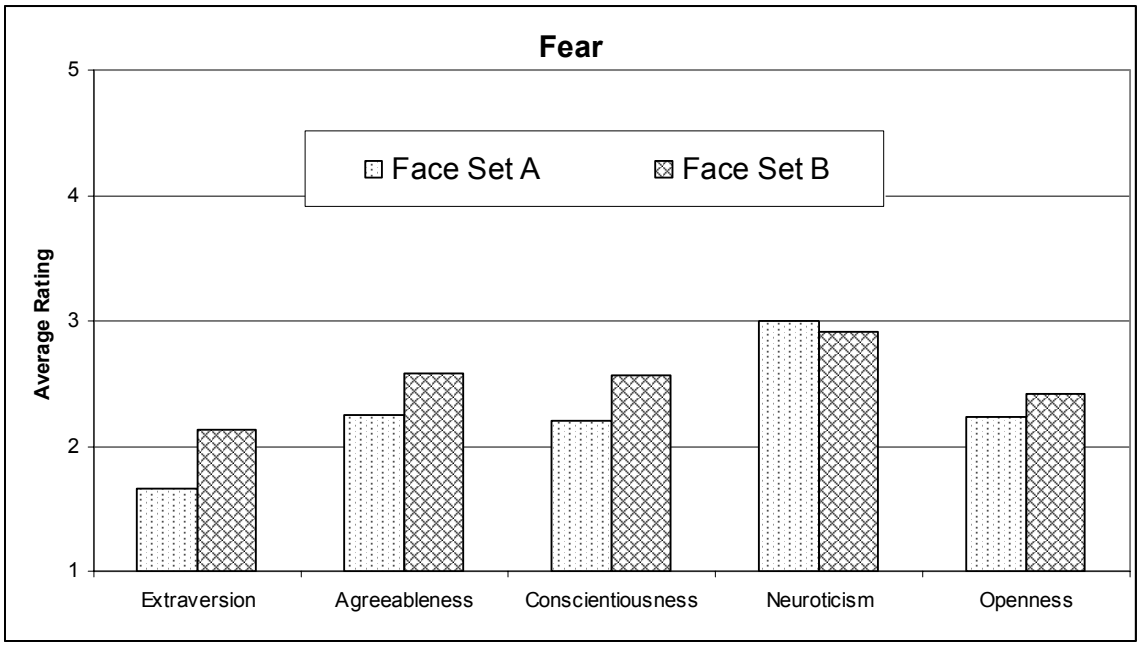


Figure 7: Average RATING for PERSONALITY TRAIT by FACE SET for the EXPRESSION Fear

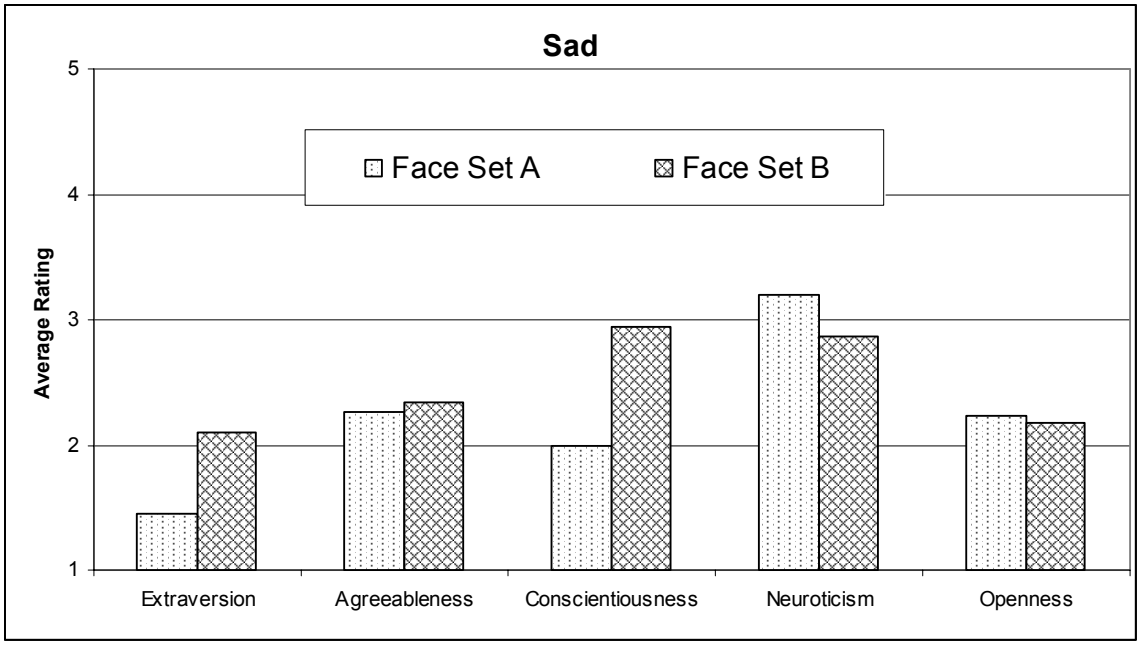


Figure 8: Average RATING for PERSONALITY TRAIT by FACE SET for the EXPRESSION Sad

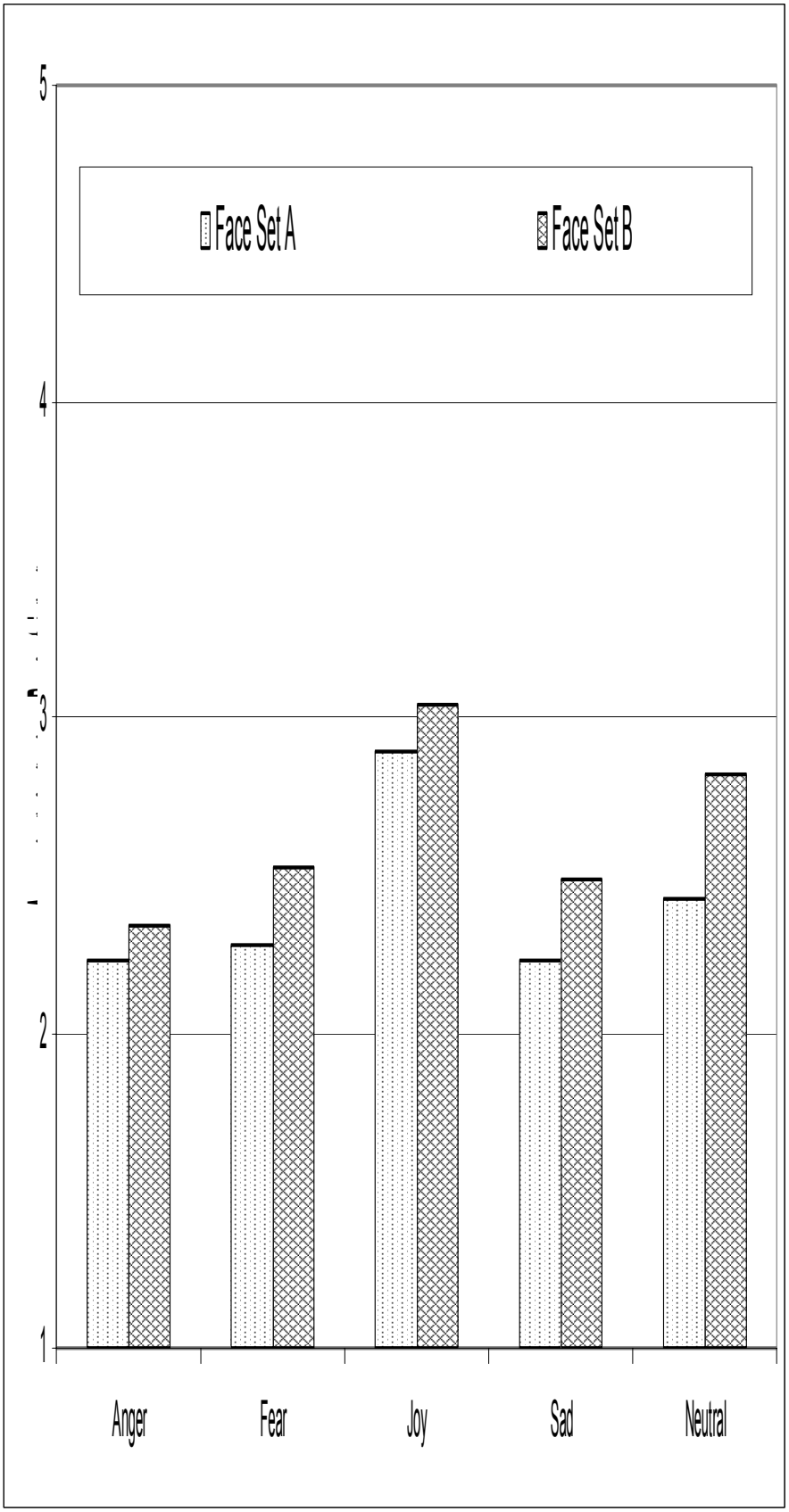


Figure 9: Average RATING for FACE SET by EXPRESSION

A second simple-effects F-test was conducted by PERSONALITY TRAIT. There was a significant interaction between FACE SET and EXPRESSION for two levels of PERSONALITY TRAIT: Conscientiousness ($F(4, 216) = 4.2545$, $p = 0.0025$) and Extraversion ($F(4, 216) = 4.4997$, $p = 0.0016$), while Agreeableness, Neuroticism and Openness did not have significant interactions. Figures 9 and 10 illustrate the significant two-way interactions. When the PERSONALITY TRAIT was Conscientiousness all five levels of EXPRESSION were rated higher for FACE SET B. The magnitude of the differences between FACE SET A and B varies depending on the emotion, with the largest difference occurring with the emotion Sad and no real difference for Anger. Similarly, when the PERSONALITY TRAIT was Extraversion, four levels of EXPRESSION were rated higher for FACE SET B. Joy was the only EXPRESSION where FACE SET A was rated higher. The pattern of differences between FACE SET A and B for Extroversion are different then the pattern for Conscientiousness. For example there is no difference in ratings for Anger in Conscientiousness, but there is a difference for Extroversion.

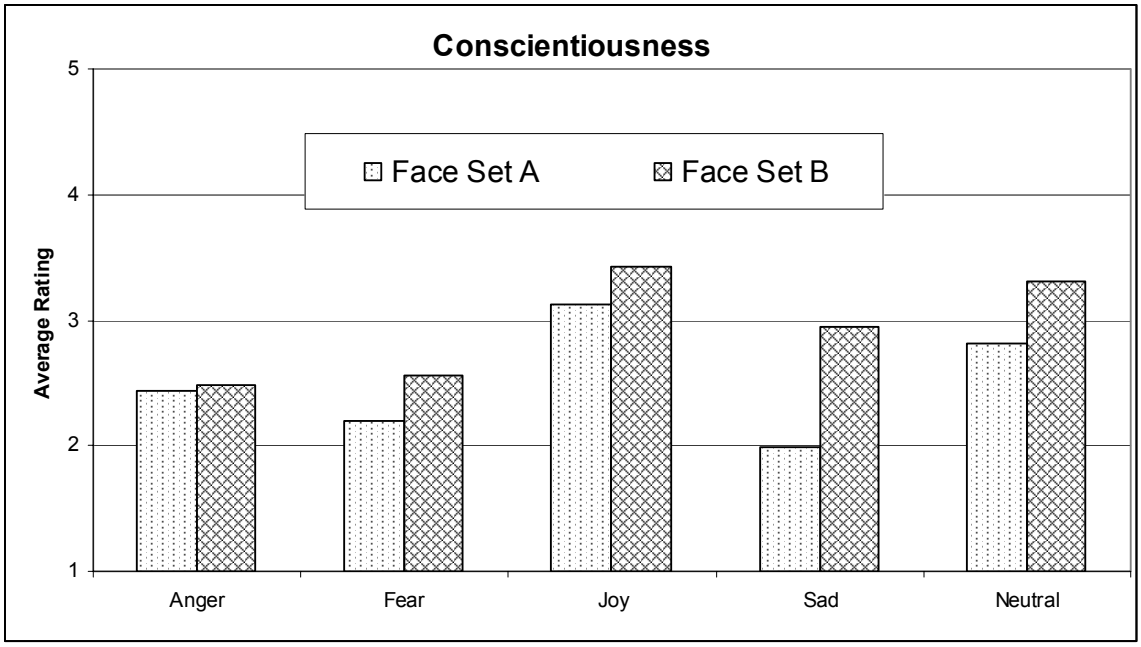


Figure 10: Average RATING for EXPRESSION by FACE SET for the PERSONALITY TRAIT Conscientiousness

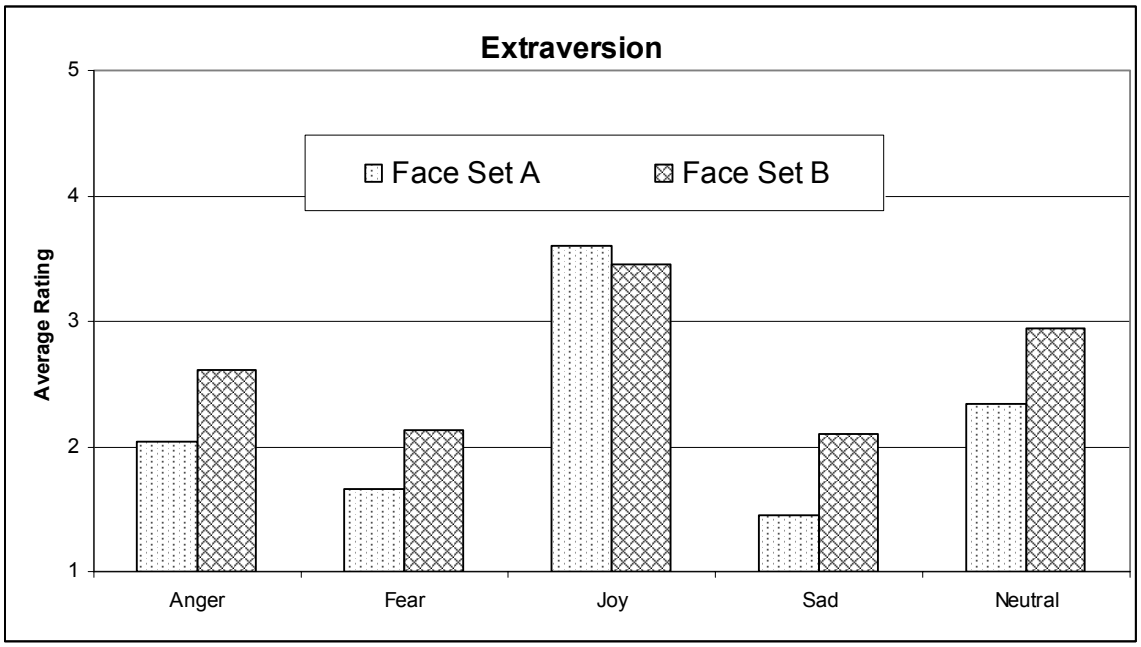


Figure 11: Average RATING for EXPRESSION by FACE SET for the PERSONALITY TRAIT Extraversion

GENDER x EXPRESSION x PERSONALITY TRAIT

There was a significant three-way interaction among GENDER, EXPRESSION and PERSONALITY TRAIT, $F (16, 864) = 2.1176$, $p = 0.0063$ as illustrated in Figure 12. A visual inspection of the graph indicates different patterns for the RATINGS based on EXPRESSION. Anger has a similar low RATING across four of the five PERSONALITY TRAITS for both Males and Females. The RATING for Joy also appears similar between Males and Females. To break down this interaction a simple-effects F-test by EXPRESSION was conducted and indicated a significant two-way interaction between GENDER and PERSONALITY TRAIT for one of the EXPRESSIONS: Sad, $F (4, 216) = 2.86161$, $p = 0.0244$ (see Figure 13). When the EXPRESSION is Sad, Male participants generally gave a higher average RATING for each PERSONALITY TRAIT with the exception of Neuroticism which had the same average RATING for both Males and Females. The average RATING for the interaction between GENDER and PERSONALITY TRAIT was not significant for the other four EXPRESSIONS.

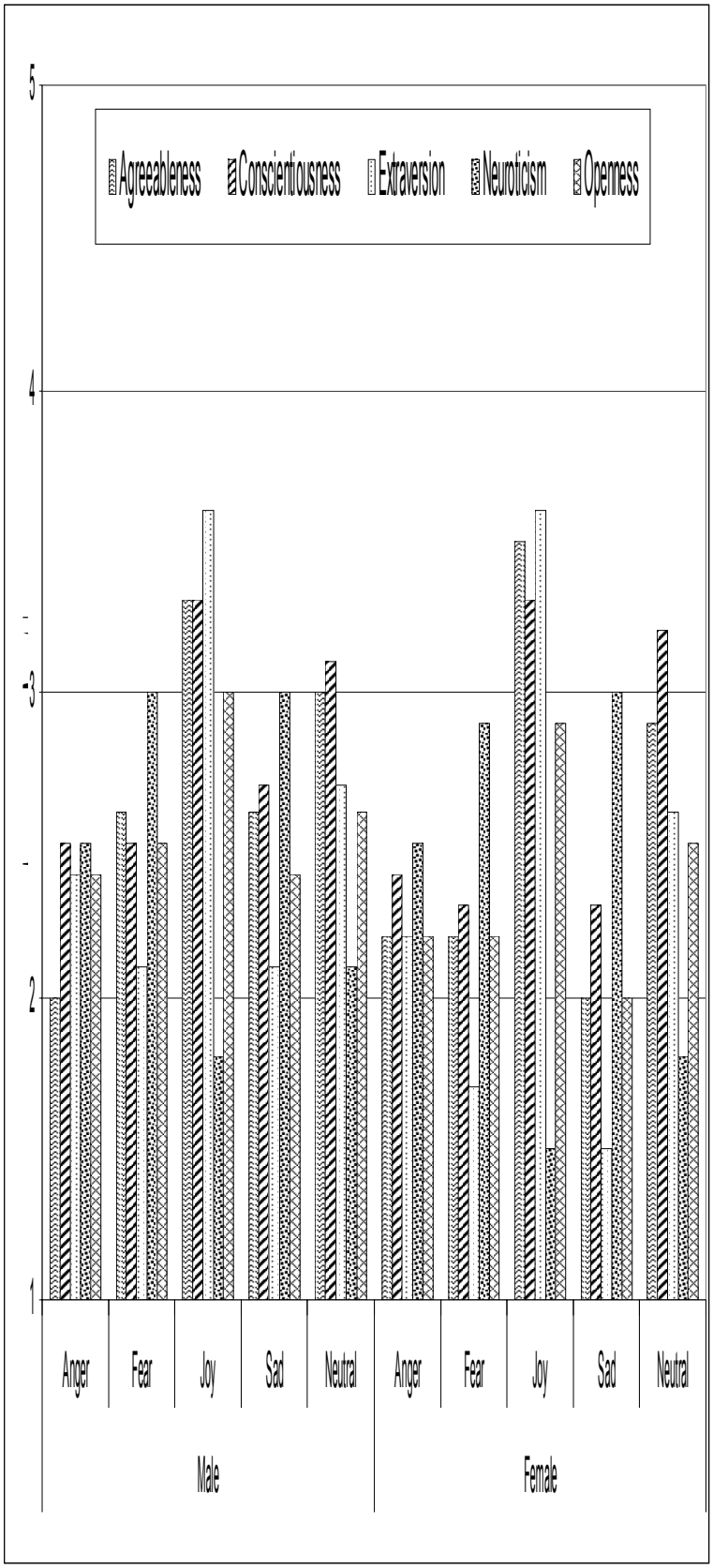


Figure 12: Average RATING for PERSONALITY TRAIT by GENDER and EXPRESSION

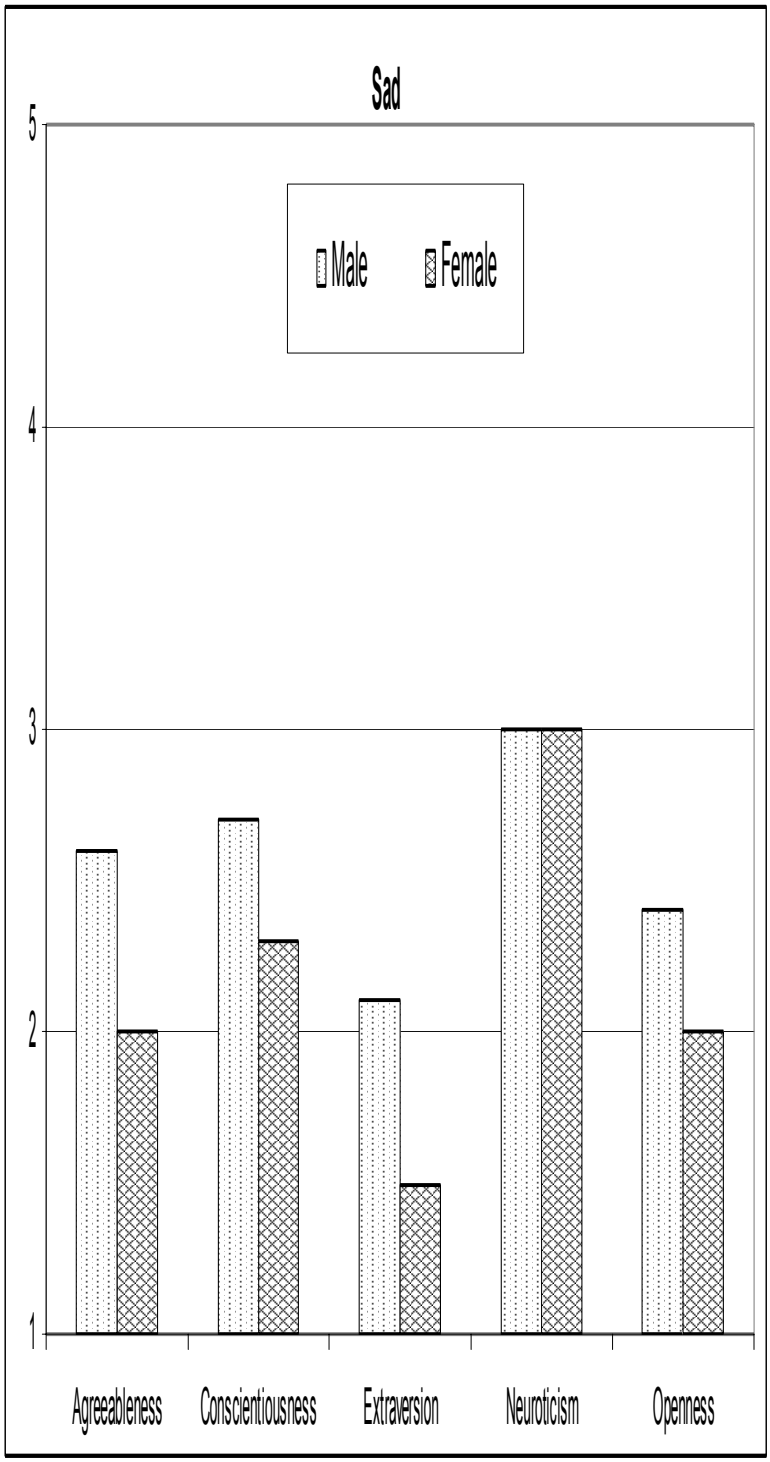


Figure 13: Average Rating for GENDER by PERSONALITY TRAIT for the EXPRESSION Sad

A simple-effects F-Test by PERSONALITY TRAIT indicated a significant interaction between GENDER and EXPRESSION for two levels of PERSONALITY TRAIT, Agreeableness and Extraversion ($F(4, 216) = 4.3632$, $p = 0.0021$ and $F(4, 216) = 3.3516$, $p = 0.0110$ respectively). As illustrated in Figure 14, when the PERSONALITY TRAIT was Agreeableness, the average RATING for Anger and Joy was rated higher among Female participants while the other EXPRESSIONS were rated higher by males. As illustrated in Figure 15, when the PERSONALITY TRAIT was Extraversion, Males gave higher average RATINGS than females except for the Expression Joy which had the same average rating for both GENDERS. In addition, the magnitude of the difference in RATINGS for the Sad EXPRESSION was greater than the difference for the other EXPRESSIONS for both the Agreeableness and Extroversion traits.

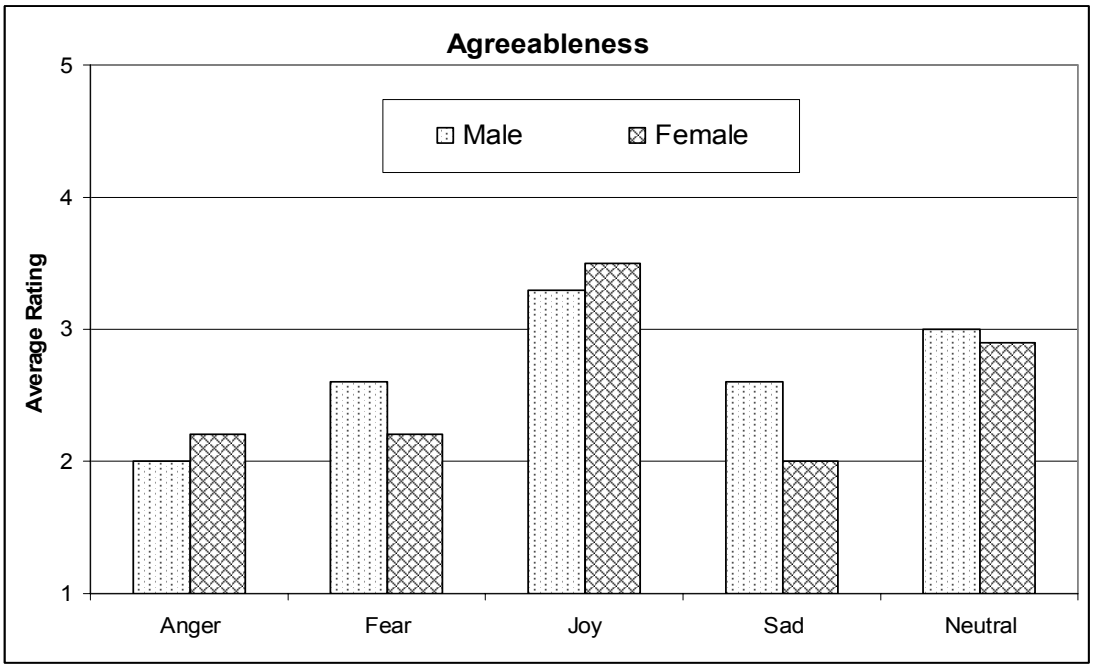


Figure 14: Average Rating for GENDER by EXPRESSION for Agreeableness

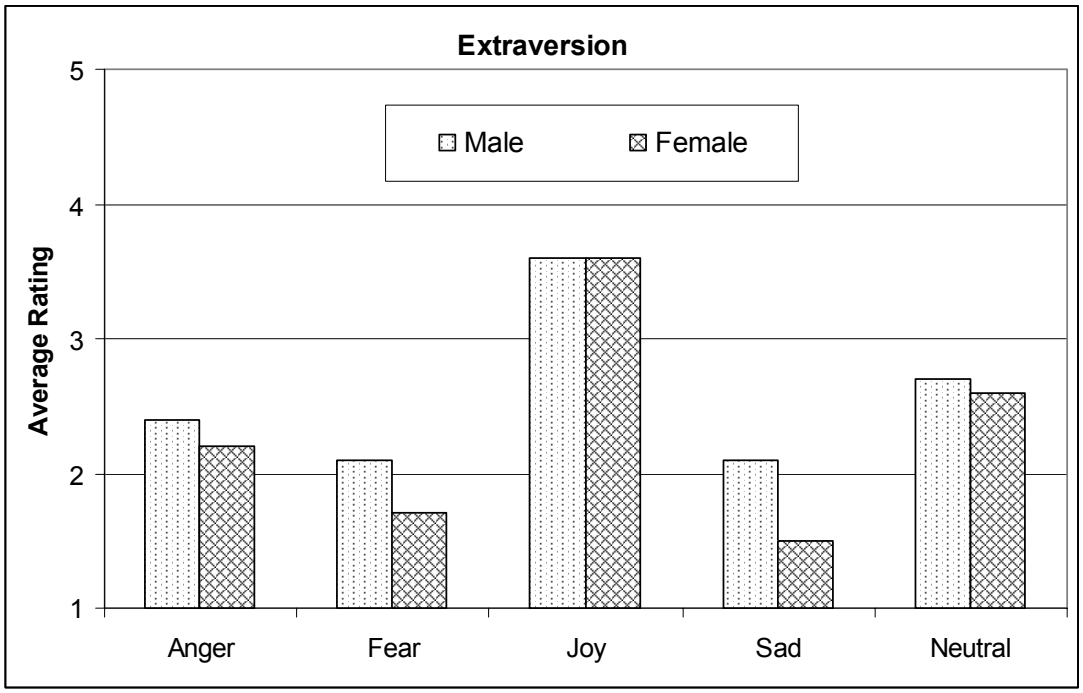


Figure 15: Average Rating for GENDER by EXPRESSION for Extraversion

7.2 Two-way Interactions

GENDER x EXPRESSION

The two-way interaction between GENDER and EXPRESSION was significant, $F(4, 216) = 3.1321$, $p = 0.0157$. Examining the simple-effects F-Test by EXPRESSION, Sad is the only EXPRESSION that showed significantly different RATINGS between the two levels of GENDER, $F(1, 54) = 11.8098$, $p = 0.0011$. As illustrated in Figure 16, the average RATING for each EXPRESSION was higher (more characteristic) for Males then for Females, but this difference was only significant for the Sad EXPRESSION.

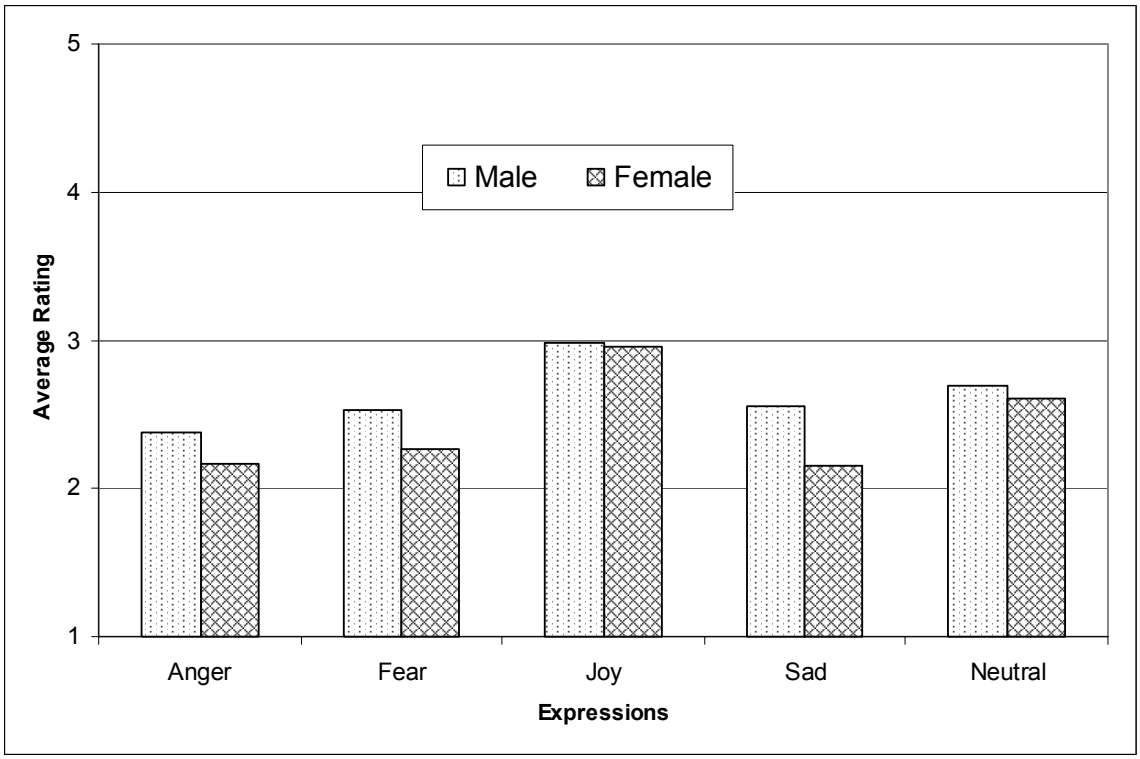


Figure 16: Average RATING for GENDER by EXPRESSION

FACE SET x PERSONALITY

The significant interaction between FACE SET and PERSONALITY TRAIT ($F(4, 216) = 6.5964$, $p < 0.0001$) is illustrated in Figure 17. The simple-effects F-Test by PERSONALITY TRAIT indicated the effect of FACE SET is significant for two PERSONALITY TRAITS, Extraversion and Conscientiousness: $F(1, 54) = 12.2157$, $p = 0.0010$ and $F(1, 54) = 6.1952$, $p = 0.0159$ respectively. As illustrated in Figure 17, the average rating for FACE SET B was significantly higher than FACE SET A for Extraversion and Conscientiousness.

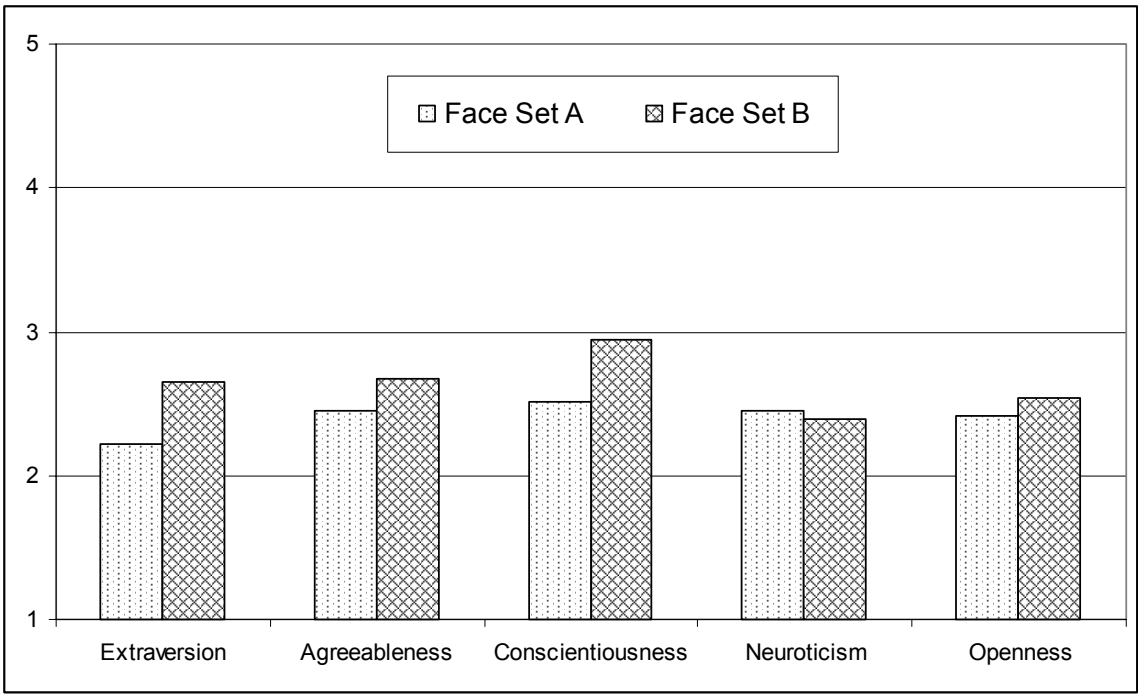


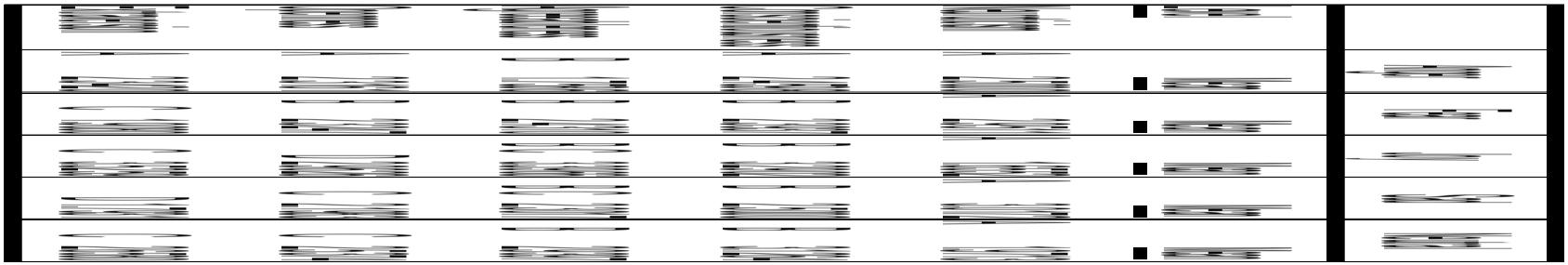
Figure 17: Average RATING for PERSONALITY TRAIT by FACE SET

EXPRESSION x PERSONALITY

The significant interaction between EXPRESSION and PERSONALITY TRAIT ($F(16, 864) = 59.8299$, $p < 0.0001$) is illustrated in Figure 18. The simple-effects F-Test by EXPRESSION indicates the effect of PERSONALITY TRAIT was significant at every level of EXPRESSION. A comparison of mean RATINGS using the Tukey-Kramer HSD, test (see Table 3 and Figure 18) shows RATINGS vary based on emotional EXPRESSION. For Anger the RATINGS were the same across all PERSONALITY TRAITS except Agreeableness. For Fear the trait Neuroticism was rated significantly higher than any of the other traits and Extroversion was rated lower than any other PERSONALITY TRAIT. For Joy, Neuroticism was rated significantly lower than any other trait followed by Openness. Extroversion and Agreeableness were rated similarly

for Joy as were Conscientiousness and Agreeableness. For the EXPRESSION Sad, Extroversion was rated lowest and Neuroticism the highest which is similar to the EXPRESSION Fear. For Neutral, Neuroticism was rated significantly lower than the other traits. Conscientiousness and Agreeableness were rated similarly for Neutral as were Extroversion and Openness. This analysis supports the idea that people perceive significant differences in personality based on emotional expression of a digital face.

Table 3: Tukey HSD comparison of mean RATINGS by PERSONALITY TRAIT



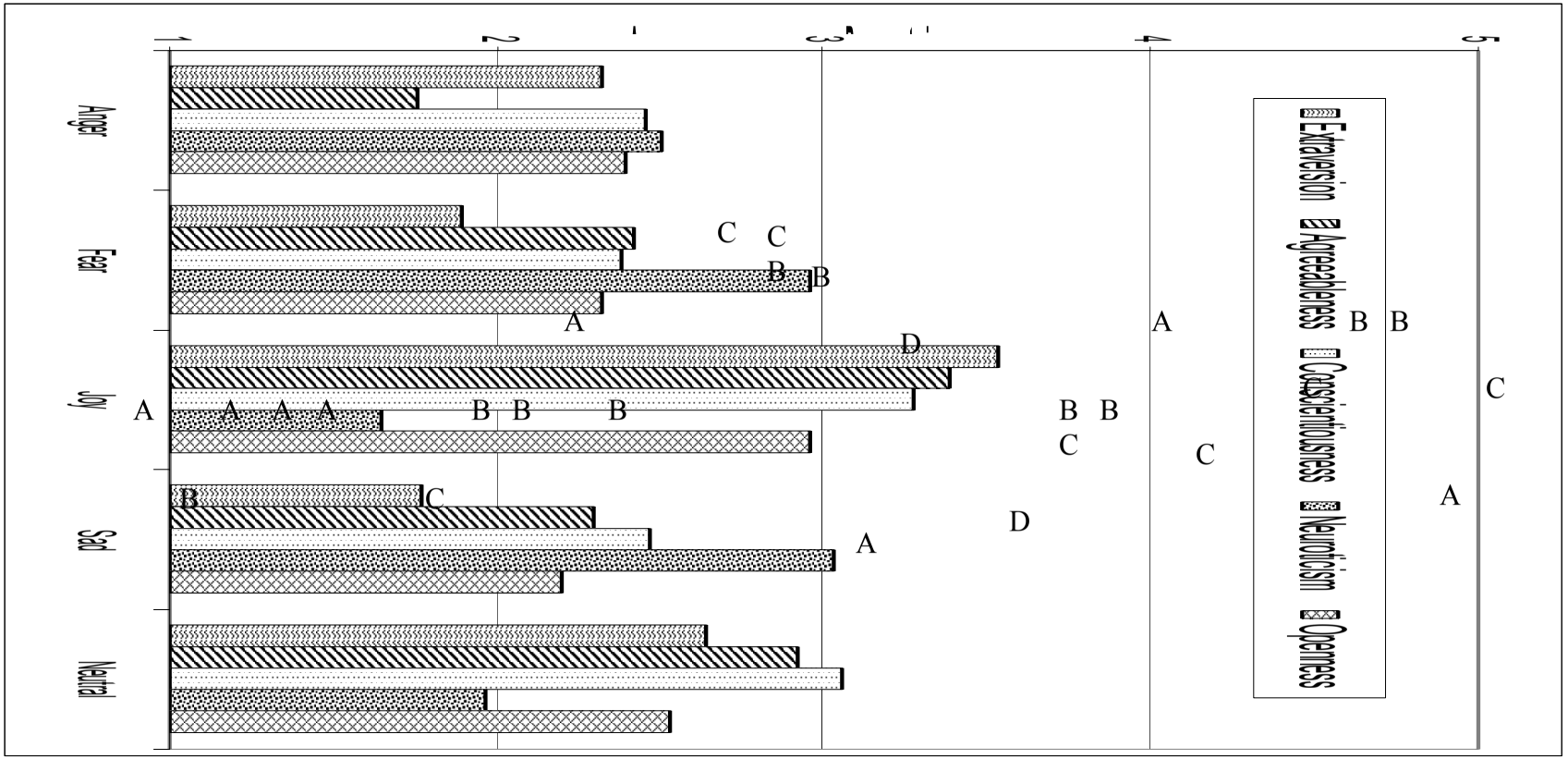


Figure 18: Average RATING for EXPRESSION by PERSONALITY TRAIT
(A, B, C and D represent the Tukey comparison of mean results)

Main Effects

The main effect of FACE SET was significant, $F(1, 54) = 4.0915$, $p = 0.0481$. The mean rating for FACE SET A ($X = 2.41$) was significantly lower than for FACE SET B ($X = 2.64$).

The main effect of EXPRESSION was significant with $F(4, 216) = 40.9827$, $p < 0.0001$. A comparison of means using the Tukey HSD test was conducted and is illustrated in Table 4. This analysis shows that the EXPRESSIONS of Joy and Neutral are significantly different than any other EXPRESSION, and Fear was significantly different from Anger. There were no other significant differences.

Table 4: Tukey HSD Comparison of Mean Ratings for EXPRESSION

Level					Mean
Joy	A				2.9661
Neutral		B			2.6466
Fear			C		2.4017
Sad			C	D	2.3621
Anger				D	2.2753

The main effect of PERSONALITY TRAIT was significant at $F(4, 216) = 11.02$, $p < 0.0001$. A comparison of means using Tukey HSD (Table 5) indicates the ratings for the PERSONALITY TRAIT Conscientiousness were significantly higher than ratings for all other traits. Ratings for Agreeableness were significantly higher than Extraversion and Neuroticism. There were no other significant differences.

Table 5: Tukey HSD Comparison of Mean Ratings for PERSONALITY TRAIT

Level				Mean
Conscientiousness	A			2.7540
Agreeableness		B		2.5655
Openness		B	C	2.4678
Extraversion			C	2.4454
Neuroticism			C	2.4190

8. PHASE I: DISCUSSION

The purpose of the first phase of this study was to determine if participants would rate static digital faces as having PERSONALITY TRAITS based on differences in emotional EXPRESSIONS. The statistical analysis showed that there are several significant interactions and main effects. To interpret these differences this discussion will focus on the six hypotheses proposed for Phase I.

First Hypothesis: *There will be no difference in the ratings based on NATIONALITY.*

Data for NATIONALITY were not analyzed due to the highly uneven number of participants in each group (48 American, 10 Indian). Previous research on Physiognomy by Zebrowitz (1997) suggests that NATIONALITY should not influence the perception of personality. Additionally Ekman, Sorenson and Frieson, (1969) showed that the five primary emotions are equally recognizable across NATIONALITY.

Second Hypothesis: *There will be no difference in ratings based on GENDER.*

Research with respect to personality traits and the Big-Five Factor model has shown the traits to be stable across gender (McCrae and Costa, 1997; Solds and Vaillant, 1999). Also previous studies related to emotional expression have found no differences between men and women in identifying the five primary emotions (Ekman, Sorenson and Frieson, 1969). Based on this research it was expected that there would be no differences in RATINGS based on GENDER. However there were some specific differences found. For the EXPRESION Sad, Males had higher average RATINGS for all five central PERSONALITY TRAITS, except Neuroticism. These differences are significant when

rating Agreeableness and Extroversion. While men and women rated Extraversion differently for Sad, they both rated that characteristic on the lower end of the scale, with RATINGS close to or below two. For Agreeableness males rated the trait higher for the Sad face. There appears to be a difference in how males and females perceived personality when a face was expressing sadness.

Third and Fifth Hypotheses: *There will be no difference in ratings based on FACE SET. There will be no difference in ratings based on EXPRESSION.*

FACE SET A and B were based on average faces and expressed similar emotions with the primary difference being that the renderings one set of face's (FACE SET B) appeared more humanlike or less cartoonish. Although this difference is not drastic it was expected that the humanlike face would be perceived differently for the same emotional EXPRESSION. It was also expected that there would be an interaction of FACE SET and EXPRESSION (Hypothesis 5).

Results indicated that there were differences related to FACE SET and EXPRESSION. There was a three-way interaction for FACE SET, EXPRESSION and PERSONALITY TRAIT. The differences between FACE SET occur for three emotional EXPRESSIONS: Anger, Fear and Sad. For Anger the only difference between RATINGS of the FACE SET are for the PERSONALITY TRAIT Extroversion. For Fear and Anger there were differences in Extroversion and Conscientiousness. When differences occur FACE SET B is generally given a higher rating than FACE SET A.

Because the primary difference between FACE SET is the type of rendering it is possible that these differences are caused by the more detailed and humanlike facial

features of stimulus B. No differences related to Joy may indicate that this EXPRESSION has very specific perceived traits regardless of whether the face has more humanlike features.

The two-way interaction between FACE SET and PERSONALITY TRAIT follows the three-way interaction closely. There was only a difference in FACE SET for two PERSONALITY TRAITS, Extroversion and Conscientiousness. The analysis of the three-way interaction indicated these differences are based on specific expressions. FACE SET B was generally rated higher than FACE SET A across the different EXPRESSIONS and PERSONALITY TRAITS, and the main effect also shows this difference.

Fourth Hypothesis: *There will be no difference in ratings based on EXPRESSION.*

It was expected that participants would rate the different emotional EXPRESSIONS differently. Previous research indicates that people differentiate between emotional EXPRESSIONS (Hagar and Ekman, 1983). The main effect of EXPRESSION was significant however not all emotions were rated differently. Sad and Anger received similar RATINGS as did Fear and Sad EXPRESSIONS. Joy and Neutral were rated differently compared to any of the other EXPRESSIONS.

This main effect was expected but not as meaningful in light of the interactions with PERSONALITY TRAIT. These interactions are discussed under hypotheses five and six.

Sixth Hypothesis: *There will be no difference in ratings between the five PERSONALITY TRAITS (Extroversion, Agreeableness, Contentiousness, Neuroticism, and Openness) based on EXPRESSION.*

It was expected that the different emotion EXPRESSIONS would result in different RATINGS of PERSONALITY TRAITS. For example the sub-trait friendliness would be related to a joyful expression, but not an angry expression. Seeing these patterns produces an indication of what PERSONALITY TRAITS are assigned to specific expressions. EXPRESSION was a variable in four of the five significant interactions (see Table 2). As expected, the RATINGS varied based on emotional EXPRESSION. Figure 18, shows the interaction of EXPRESSION by PERSONALITY TRAIT, which illustrated the overall summary of the patterns.

When the EXPRESSION was Joy, the average participant RATINGS for Extroversion, Agreeableness and Conscientiousness were above 3, Openness 3 (Partially Characteristic) and Neuroticism was rated 1.5. This pattern is consistent across FACE SET.

However when the EXPRESSION is Anger, the average is no more than 2.5, with Agreeableness being rated significantly lower. The EXPRESSION of Anger leads to the impression of not being agreeable which seems to be a likely characterization.

For the EXPRESSIONS of Sad and Fear the average RATINGS for Neuroticism were near 3 and significantly higher than the RATINGS for any of the other traits. Sad and Fear EXPRESSIONS also have a significantly lower RATING for the PERSONALITY TRAIT of Extroversion. This leads to the impression that both Sad and Fearful faces give the sense of being neurotic while lacking the characteristics of an

extrovert. Considering the sub-traits (Anxiety, Anger, Depression, Self-Conscientiousness, Immoderation and Vulnerability) that are incorporated in the Neurotic PERSONALITY TRAIT this characterization seems likely.

For Neutral the highest RATINGS were for Agreeableness and Conscientiousness at approximately 3, while Neuroticism was rated significantly lower than the other PERSONALITY TRAITS at less than 2. These RATINGS lead to the impression that a Neutral EXPRESSION is perceived to have a personality similar to a Joyful face, but with lower average ratings.

Additional Observations

It is interesting to note that the average RATINGS tended to fall between 1.5 and 3, with some averages higher than 3 for the emotional EXPRESSIONS of Joy and Neutral (See Figure 18). Lower ratings closer to 1 indicate a trait that is not perceived and RATINGS around 3 indicate that a trait is at least partially characteristic.

Because each central PERSONALITY TRAIT is determined by averaging the data from six sub-traits, it is possible that the averaging causes the score to be low while masking certain sub-traits that were rated quite high. One example of this is the PERSONALIY TRAIT of Extroversion which includes the sub-trait Cheerfulness and is easily identified in a Joyful expression. However, one of the other sub-traits for Extroversion is Assertiveness, which may not have the same logical connection to a Joyful EXPRESSION. The result of averaging these characteristics masks the individual importance of sub-traits. Further analysis based on sub-traits is necessary to determine perceptions of sub-traits that are more strongly produced by facial EXPRESSIONS.

9. STIMULUS SELECTION FOR PHASE II

The three stimuli used for Phase II were selected from the 10 faces evaluated during Phase I. Selection was based on perceived personality as determined from the average RATINGS for each face in Phase I. The 30 sub-trait RATINGS were grouped and averaged into the Big-Five Factor PERSONALITY TRAITS and then compared to the personality profile of the "ideal collaborative partner." The "ideal collaborative partner" as described by Prabhala and Gallimore (2005) is rated higher for extraversion, agreeableness and conscientiousness, with low neuroticism and moderate openness.

The three face stimuli selected for Phase II were the two faces whose personalities most closely matched the "ideal collaborative partner" and the one stimulus that was least related to the "ideal collaborative partner" personality. The two face stimuli whose perceived personality most closely matched that of the "ideal collaborative partner" are faces A3 and B3, the two faces with a Joyful EXPRESSION. As illustrated in Figure 5 the personality profile of both faces is nearly a perfect match as these faces have higher ratings for extraversion, agreeableness and conscientiousness, with a low rating for neuroticism and moderate rating for openness. The hierarchy cluster analysis shown in Figure 19 shows that faces A3 and B3 are more closely related than any of the other face stimuli. The face with a personality that was least similar to the "ideal collaborative partner" was face A4, the sad face in set A. As shown in Figure 5 the ratings for the PERSONALITY TRAITS of face A4 are a proportional inverse of faces A3 and B3. This dissimilar relationship is also supported by the cluster hierarchy (Figure 19). Therefore, the three faces selected for phase II were A3, B3 and A4 (Figure 20).

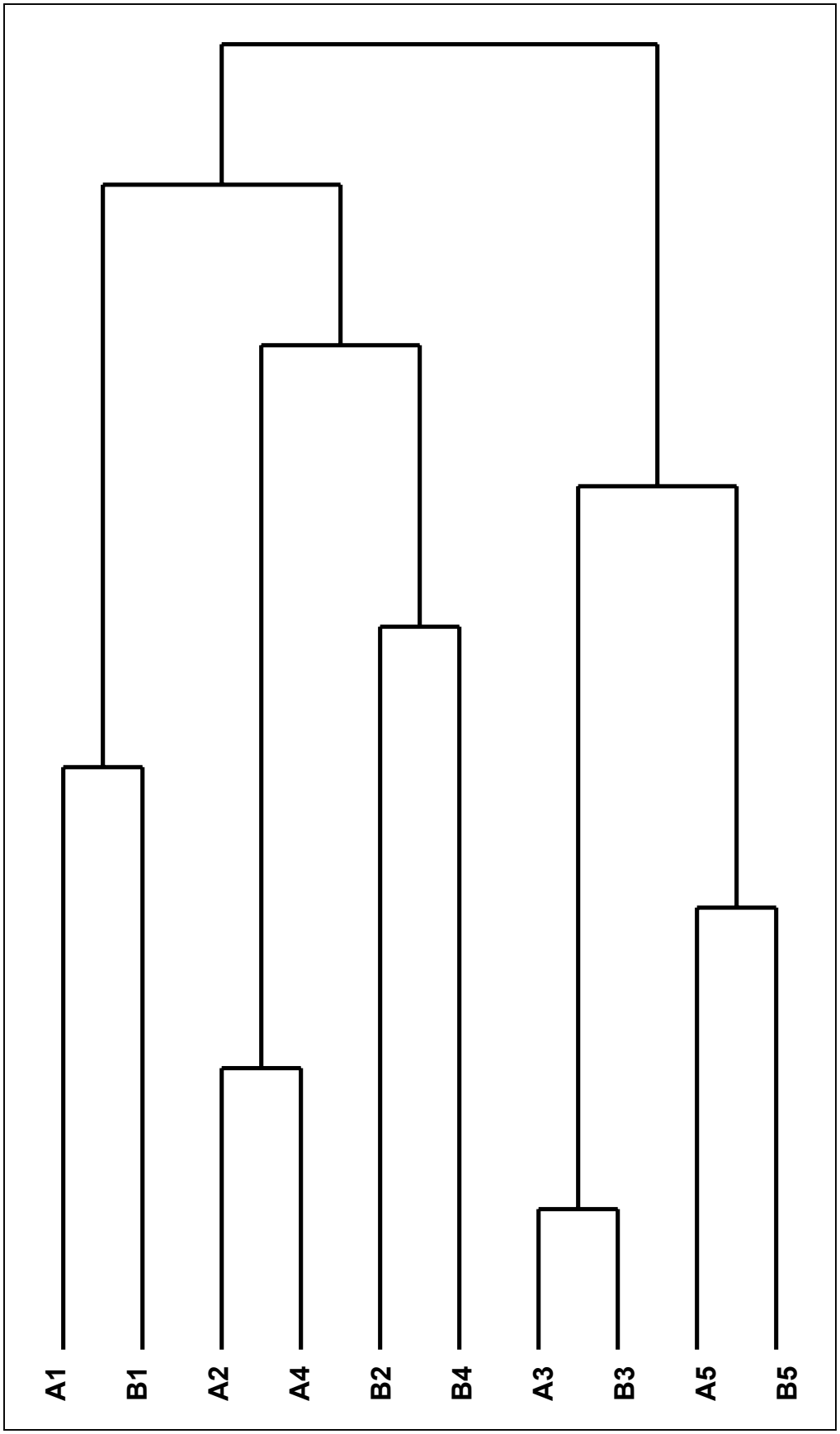


Figure 19: Hierarchal cluster analysis of the 30 individual sub-trait RATINGS for the digital faces

7. PHASE II: HYPOTHESES

The second phase of the study will examine whether the projected personality of a digital face influences performance in a human-machine collaborative task. The following six hypotheses will be tested:

Table 6: Phase II - Hypotheses and Expectations

	Null Hypotheses	Expectation
1.	There will be no difference in Task Score and Suggestion Use based on participant **Nationality**.	Fail to Reject Hypothesis
2.	There will be no difference in Task Score and Suggestion Use based on participant **Gender**.	Reject Hypothesis
3.	There will be no difference in Task Score based on **Face** stimuli.	Reject Hypothesis
4.	There will be no difference in Suggestion Use based on **Face** stimuli.	Reject Hypothesis
5.	There will be no difference in Task Score based on the level of **suggestion accuracy**.	Reject Hypothesis
6.	There will be no difference in Suggestion Use based on the level of **suggestion accuracy**.	Reject Hypothesis

Testing the hypotheses listed in Table 6 will provide insight into how the facial representation and suggestion accuracy affect the collaborative interaction. In addition we can distinguish differences based on GENDER and NATIONALITY.

11. PHASE II: METHOD

11.1 Experimental Design

The experiment for the second phase of research was designed to present a series of three collaborative tasks in which the subject would work with a computer agent who provided suggestions. Each of the three tasks represented a single trial. Two dependent variables, task score and number of suggestions used, were recorded for each trial. Participants were paired with a different stimulus for each of the three tasks while a single level of suggestion accuracy was assigned across all three tasks. The order of the tasks, the stimulus assigned to each task, and the level of suggestion accuracy were ordered using a factorial matrix to maximize the usefulness of the data and minimize order effects. The experimental design for Phase II is a 3 x 3 x 2 x 2 mixed factorial design. The between-subject variables are three levels of suggestion accuracy (good, moderate, poor), nationality (American, Indian) and gender (male, female). The within-subject variables are three face stimuli (A3, B3, A4 (Figure 20)). A full description of each scenario is provided in Appendices E, F and G .

11.2 Subjects

Thirty-six subjects from Wright State University participated as volunteers for the second phase of this study. All subjects were engineering students blocked into four groups (American-Males, American-Females, Indian-Males, Indian-Females) consisting of nine subjects in each group.

11.3 Stimuli

The three stimuli used for Phase II were selected from the ten faces evaluated in the first phase of this study. These three faces were selected based on perceived personality as determined from their ratings in Phase I. Characteristic ratings were grouped into the Big-Five Factor personality TRAITS and then compared to the personality of the "ideal collaborative partner" as described by Prabhala and Gallimore (2005). The two faces whose perceived personality most closely matched that of the "ideal" personality and a third face with a personality that was least similar to the "ideal" partner were used. The faces selected were A3, B3 and A4 (Figure 20).

Figure 20: Stimuli used for Phase II. A3-Joy, B3-Joy, A4-Sad

11.4 Apparatus

For the second phase of the study subjects were seated in the Cacioppo Laboratory at Wright State University. To complete the activities in Phase II, subjects used a desktop computer with a 19-inch LCD monitor and a two button mouse.

Participants first filled out the IPIP NEO-Personality Index survey hosted at www.personal.psu.edu/faculty/j/5/j5j/IPIP/ipipneo120.htm. This required participants to use Microsoft's Internet Explorer and answer questions by selecting radio buttons with the mouse.

The program used for the second activity in Phase II was written in visual basic.NET using the Microsoft.NET Developer. The program was used to provide the series of interactive tasks for Phase II, including algorithms to provide the three levels of suggestion accuracy, task scoring and tracking of suggestions used. Participants completed these tasks using the mouse to select items from a dropdown box.

11.5 Procedure

After signing the consent form, subjects for the second phase of research completed two activities. First, complete the short form IPIP NEO-Personality Index, and then complete the series of three collaborative tasks.

The short form IPIP NEO-Personality Index is a personality survey consisting of 120 statements which participants rates according to how accurately each statement describes their feelings or actions. The result of this survey is formatted according to the Big Five Factor personality model, providing scores for each of the five primary personality TRAITS and individual scores for the six sub-TRAITS that makeup each trait. After completing the personality survey, subjects begin the second activity.

The second activity was a series of three collaborative tasks in which the participant worked with the computer agent to rank a list of items in a survival scenario. Each task began with an introduction screen where the participant was shown a display

with only a picture of the stimulus and a button to begin the scenario (Figure 21). At this point participants were given verbal instructions stating that this face represented the computer agent they would be working with during the scenario and that this computer agent would be making two suggestions for each ranking selection and that suggestions would update each time one of the suggestions was used. Information about the agent's expertise or other TRAITS was specifically excluded.

Figure 21: Phase II introduction screen

After selecting the button to begin, the subject was presented with a brief written description of a "survival scenario" including a list of objects that were available to aid in their survival. In addition there was a picture of the stimulus face, a text box with two

suggestions and a dropdown box for selecting their first choice (Figure 22). The participants were asked to read the scenario and list of items. Their task was to rank the items in order of value, selecting the most useful item first. Participants were given brief verbal instructions and shown how to select their first choice. Before selecting the first item, the computer agent, represented by the stimulus face provided two suggested items in the text box (Figure 22). The suggestions had varying degrees of accuracy: good (best two choices), moderate (two choices from the middle of the list) and poor (worst two choices). The level of suggestion accuracy was not known by the participant and remained constant throughout all three trials for a given subject. Once they selected their first choice, an "accept" button appeared next to the selection box.

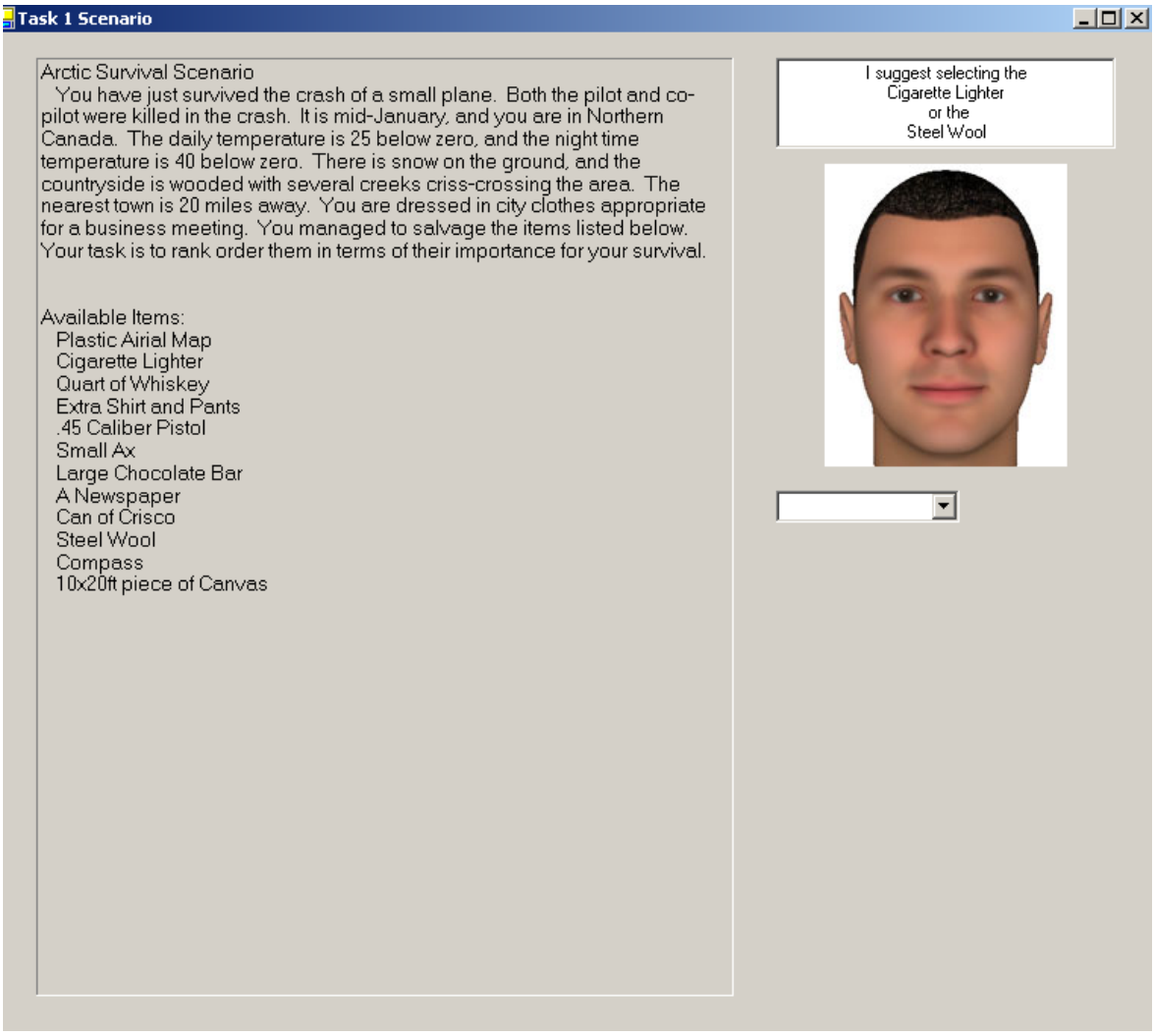


Figure 22: Phase II survival task initial view

At this point subjects could still change their selection. Pressing the "accept" button locked in the first choice and a dropdown box appeared to select their second choice. The suggestions were updated as they were used and remained constant when not used. Items that had already been selected no longer appeared on the dropdown list (Figure 23). This process was repeated until the top 10 items were ranked. After completing the first scenario subjects proceded with the second and third scenarios

starting with the introduction screen and then repeating the same steps followed in the first task.

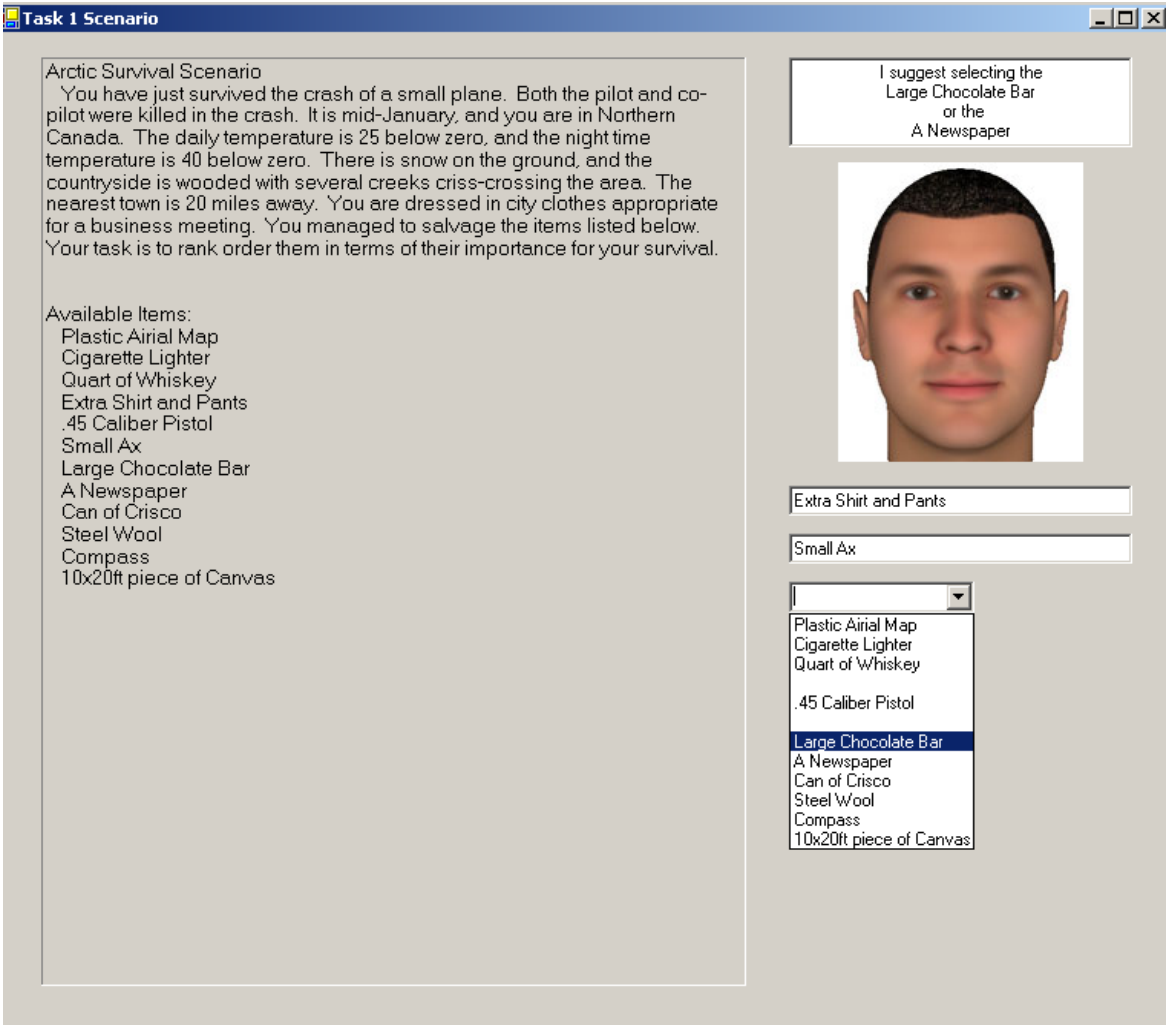


Figure 23: Phase II survival task in progress

After completing the three tasks subjects were shown their scores and thanked for their participation.

11.6 Dependent Variables

For each of the three tasks, the total score and number of agent suggestions used by the participant were recorded as the dependent variables for each trial. Task score was calculated by summing the differences between the rank of each item as selected by the participant and the rank of the item as assigned by experts. Therefore a perfect score would be zero while the worst possible score is sixty. Expert rankings were provided with each scenario description. The 36 subjects participating in three trials represent the 108 data points for each of the two dependant variables

12. PHASE II: RESULTS

Subject data were analyzed using JMP IN 5.1 statistical software package. Six data points (three for each dependant variable, SCORE and SUGGESTION USE) were collected from each of the 36 subjects. The mean task SCORE was 28.59, where zero is perfect and 60 is the worst possible score. The mean SUGGESTION USE was 4.63 out of a possible 10 opportunities. An ANOVA was run for both SCORE (Table 7) and SUGGESTION USE (Table 8). Using a significance criterion of 0.05, the results showed significance for some main effects, but there were no significant interactions for either dependent variable.

Task Score

Table 7 shows that only two of the main effects are significant for task SCORE, with no significant interactions. The first main effect, NATIONALITY is significant, $F(1, 24) = 8.6465$, $p = 0.0071$. As illustrated in Figure 24, Indian participants performed worse (scored higher, $X_{Indian} = 31.56$) than American participants ($X_{American} = 25.63$).

The other significant effect was SUGGESTION ACCURACY, $F(2, 24) = 4.8890$, $p = 0.0166$. Using the Tukey-Kramer HSD test to compare the mean SCORES, we see that participants scored significantly better (lower) when receiving "Good" suggestions ($X_{Good} = 24.14$) while there was no significant difference between participants who received "Moderate" or "Poor" suggestion accuracy levels ($X_{Moderate} = 30.69$, $X_{Poor} = 30.94$). These relationships are illustrated in Figure 25. This result is expected. When participants use better suggestions their scores should improve.

Table 7: ANOVA results for task SCORE

Between Subject	df	SS	MS	F Ratio	Prob > F
Nationality	1	948.1480	948.1480	8.6465	0.0071
Gender	1	13.3704	13.3704	0.1219	0.7300
Sugst Level	2	1072.2400	536.1200	4.8890	0.0166
Nationality*Gender	1	208.3330	208.3330	1.8999	0.1808
Nationality*Sugst Level	2	15.0185	7.5093	0.0685	0.9340
Gender*Sugst Level	2	436.4630	218.2310	1.9901	0.1586
Nationality*Gender*Sugst Level	2	26.0556	13.0278	0.1188	0.8885
Subject[Nationality,Gender,Sugst Level]	24	2631.7800	109.6570		
Within Subject					
Face	2	254.2960	127.1480	1.1934	0.3120
Nationality*Face	2	79.6296	39.8148	0.3737	0.6902
Gender*Face	2	583.1850	291.5930	2.7368	0.0749
Face*Sugst Level	4	58.2037	14.5509	0.1366	0.9680
Nationality*Gender*Face	2	180.6670	90.3333	0.8478	0.4347
Nationality*Face*Sugst Level	4	749.8700	187.4680	1.7595	0.1525
Gender*Face*Sugst Level	4	76.6481	19.1620	0.1798	0.9477
Nationality*Gender*Face*Sugst Level	4	129.9440	32.4861	0.3049	0.8733
Subject*Face[Nationality,Gender,Sugst Level]	48	5114.2200	106.5460		

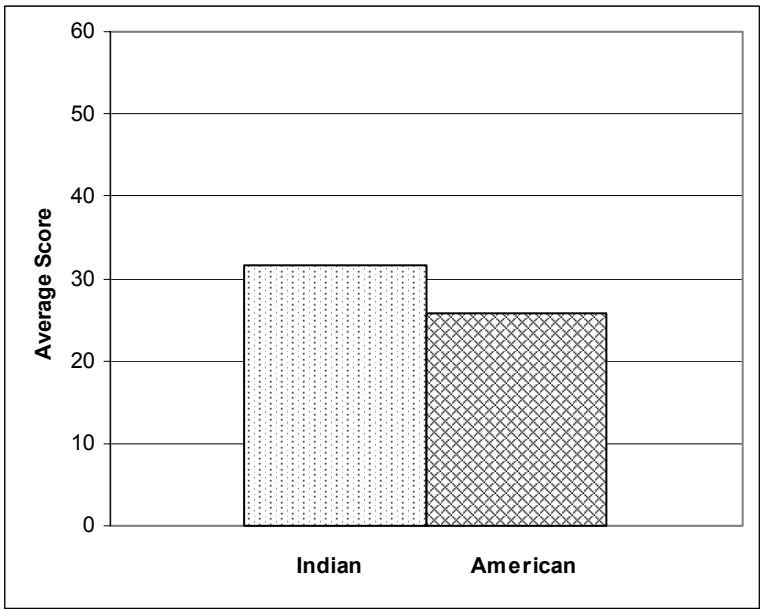


Figure 24: Average SCORE by NATIONALITY

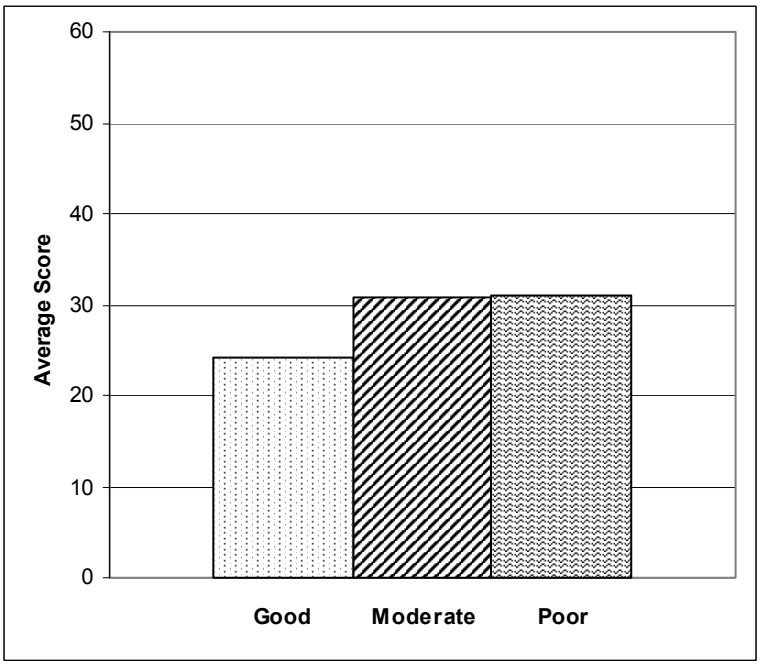


Figure 25: Average SCORE by SUGGESTION ACCURACY

Suggestion Use

Table 8 shows that only two of the main effects for SUGGESTION USE were significant: GENDER ($F(1, 24) = 7.8904$, $p = 0.0097$) and SUGGESTION ACCURACY ($F(2, 24) = 8.4490$, $p = 0.0017$). As illustrated in Figure 26, Male participants used fewer suggestions ($X_{Male} = 3.963$) than Female participants ($X_{Female} = 5.296$).

For SUGGESTION ACCURACY the Tukey-Kramer HSD was conducted to compare the mean SUGGESTION USE. The average number of suggestions used was greater when participants received "Good" SUGGESTION ACCURACY ($X_{Good} = 5.92$). Although there appears to be a trend of increasing SUGGESTION USE as SUGGESTION ACCURACY increases (Figure 27), the Tukey test shows that the number of suggestions used is not significantly different when comparing "Moderate" to "Poor" levels of SUGGESTION ACCURACY ($X_{Moderate} = 4.42$, $X_{Poor} = 3.56$).

Table 8: ANOVA results for SUGGESTION USE

Between Subject	**DF**	**SS**	**MS**	**F Ratio**	**Prob > F**
Nationality	1	5.3333	5.3333	0.8767	0.3584
Gender	1	48.0000	48.0000	7.8904	0.0097
Sugst Level	2	102.7960	51.3981	8.4490	0.0017
Nationality*Gender	1	7.2593	7.2593	1.1933	0.2855
Nationality*Sugst Level	2	6.5000	3.2500	0.5342	0.5929
Gender*Sugst Level	2	1.7222	0.8611	0.1416	0.8687
Nationality*Gender*Sugst Level	2	6.2407	3.1204	0.5129	0.6052
Subject[Nationality,Gender,Sugst Level]	24	146.0000	6.0833		
Within Subject					
Face	2	8.9074	4.4537	1.1833	0.3150
Nationality*Face	2	3.1667	1.5833	0.4207	0.6590
Gender*Face	2	9.5000	4.7500	1.2620	0.2923
Face*Sugst Level	4	8.9815	2.2454	0.5966	0.6669
Nationality*Gender*Face	2	3.6852	1.8426	0.4895	0.6159
Nationality*Face*Sugst Level	4	9.8333	2.4583	0.6531	0.6275
Gender*Face*Sugst Level	4	14.2778	3.5694	0.9483	0.4444
Nationality*Gender*Face*Sugst Level	4	2.3148	0.5787	0.1538	0.9604
Subject*Face[Nationality,Gender,Sugst Level]	48	180.6670	3.7639		

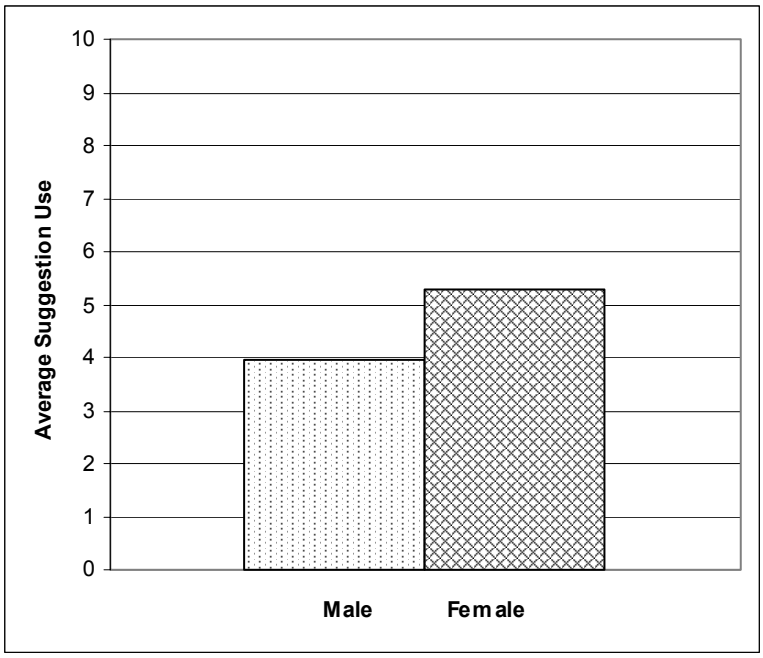


Figure 26: Average SUGGESTION USE by GENDER

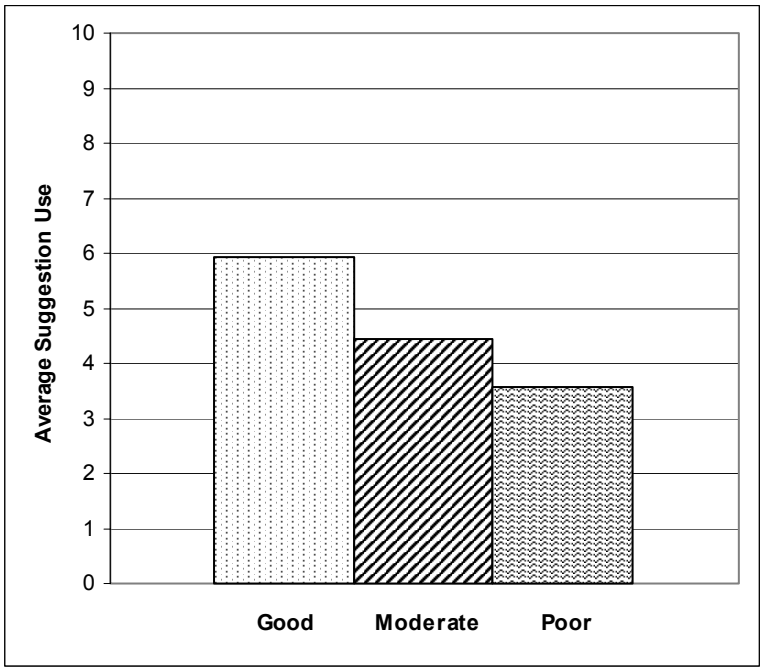


Figure 27: Average SUGGESTION USE by SUGGESTION ACCURACY

13. PHASE II: DISCUSSION

The purpose of the second experiment was to examine if the projected personality of a digital face influenced performance in a human-machine collaborative task. The statistical analysis showed that there were a few significant main effects, but no significant interactions. To explore the meaning of these results this discussion will focus on the hypotheses proposed for Phase II.

First Hypothesis: *There will be no difference in task SCORE and SUGGESTION USE based on participant NATIONALITY.*

The ANOVA results for SCORE showed that NATIONALITY had a significant influence on task SCORE. As illustrated in Figure 24, Indian participants had a higher mean SCORE. This higher SCORE represents poorer performance. The difference in average SCORE may be caused by the type of task used for this research. Although data were not collected, this difference in task SCORE may be attributed to familiarity with the type of task used for this experiment. Through interaction with the participants, it was evident that Americans were familiar with the task of ranking items in a survival scenario while it was a new experience for many Indian participants.

In contrast, NATIONALITY did not significantly influence SUGGESTION USE. Given that Indian participants may be more unfamiliar with the task, one might expect they would use more suggestions. Therefore it is difficult to pinpoint the real difference in performance between NATIONALITY.

Second Hypothesis: *There will be no difference in task SCORE and SUGGESTION USE based on participant GENDER.*

The results indicated no significant difference between Male and Female participants on TASK SCORE. However there was a difference for SUGGESTION USE with Females using more suggestions. Perhaps, women are more open to suggestions, as often indicated by stereotypes.

Third Hypothesis: *There will be no difference in task SCORE based on the FACE stimuli.*

The ANOVA for task SCORE showed that the stimulus FACE did not significantly influence task SCORE. Therefore, we fail to reject this hypothesis, indicating that the FACE of the computer agent did not influence task SCORE, even though subjects assigned PERSONALITY TRAITS to static facial EXPRESSIONS in Phase I. The most likely explanation for this finding is that the subjects were not required to directly interact with the computer agent to perform the task. Also, while subjects in Phase I rated the faces to have PERSONALITY TRAITS when asked, there was no direct link between the personality and the suggestions provided or the method for presenting the suggestions. Providing more interaction with an agent in which personality can be more easily coupled with suggestions for a more collaborative experience may produce different results.

Fourth Hypothesis: *There will be no difference in SUGGESTION USE based on the FACE stimuli.*

The ANOVA for SUGGESTION USE shows that the stimulus FACE as a main effect did not have a significant influence on the number of suggestions used by the participant. Therefore, we fail to reject the hypothesis and conclude that the FACE of the collaborative partner did not influence SUGGESTION USE. Similar to the previous hypothesis, the results of this hypothesis may be a result of limited interaction between the participant and the computer agent or may show that the stimulus FACE does not have a significant impact on SUGGESTION USE.

Fifth Hypothesis: *There will be no difference in task SCORE based on the level of SUGGESTION ACCURACY.*

The ANOVA for task SCORE shows that the level of SUGGESTION ACCURACY (Good, Moderate, Poor) does influence participant SCORE. Participants SCORE was significantly better when receiving "Good" SUGGESTION ACCURACY no significant difference in SCORES occurred when participants received "Moderate" or "Poor" SUGGESTION ACCURACY. Apparently participants were focused on the task and recognized when suggestions were not Good.

Sixth Hypothesis: *There will be no difference in SUGGESTION USE based on the level of SUGGESTION ACCURACY.*

Similar to the findings for task SCORE, when participants received Good suggestions they used them. When the SUGGESTION ACCURACY was Moderate or

Poor, the number of SUGGESTIONS USED was lower and not significantly different. As expected, participants appeared to consider the quality of the suggestions and used them if they were good and ignored less desirable suggestions.

14. FUTURE RESEARCH

The findings in Phase I draw attention to the complexity of user perception of facial representations. This experiment took one of the first steps in building a foundation to understand the visual aspects of a computer agent in relation to emotional expression and perceived personality. It is necessary to continue building on this foundation and to design computer agents with personality, so the interaction between computers and human operators can continue to evolve and improve.

The lack of findings for Phase II was likely due to lack of required interaction with the computer agent during the task. Despite the lack of interaction, the results emphasize the need to attend to differences in participant gender and nationality. Additionally, the quality of information provided by the agent impacts how users make use of the data. The results and experience gained from both phases of this study evoke several possible research opportunities for the future.

Expression and Personality

This study looked at how two similar facial representations with emotional expressions influenced user perception of agent personality. This line of research is currently very limited for human-human interaction and is nonexistent for human-computer interaction. The first step for continuing this research should further examine how emotional expression influences the perception of personality in computer agents. It is essential to have a better understanding of how expressions influence the perception of personality in order to design computer agents with specific personalities. To gain this understanding a study should examine a wide variety of faces. These faces should vary in

numerous dimensions including nationality, gender, head shape, hair style, realism (lifelike), detailed, whimsical, simple, and abstraction, etc. The results should focus on differences in perception between emotional expression and facial features. This would help determine which expressions are perceived more consistently, the impact of facial features (shape, color, detail, etc.) and any interactions that may exist. Once these interactions are understood, further steps can be taken for applying expressions to the design of computer agents.

Actions and Motion

The next step in building a computer agent that provides more natural interaction would be to expand the use of emotional expression to project personality by adding motion (changes in expression) and action to create more realism and stronger perceived personality. This research could benefit from the use of tools in the area of affective computing and augmented cognition, to enable the computer agent to react to a user's actions. Research should examine which motions and actions strengthen the perception of personality. The motions and actions would be a form of feedback and would need to be examined to determine how they influence the perception of personality. Two important factors to consider include magnitude and frequency of the motions in relation to the actions and tasks of the user. Results from systematically studying facial motions would provide a foundation for building more complex interactions between the computer and the human operator.

Multi-modal interactions

To provide a human-computer interface that represents how people work with each other, it is necessary to provide a multi-modal interface. This interface would not only account for changes in facial expression, but also include other aspects of human interaction. Additional modes of interaction include gestures, language, voice and touch. To incorporate these interactions would require complex tracking of the operator's state and actions. Prabhala and Gallimore (2005c) have begun creating a taxonomy of actions, language, and behaviors that lead to perceptions of personality. Nonverbal behaviors including facial expression are important for this taxonomy. Prabhala and Gallimore have developed a multi-modal interface that currently includes voice, visual and tactile output from the computer agent; however, facial expressions are not currently included. The complex interaction among all types of verbal and non-verbal communication is important to the development of computer agents with personality. The current project provides important input to the development of the taxonomy and the multi-modal interface.

Improve User Performance

Although studying perceptions in perceived personality may be interesting, the end goal should always focus on how these perceptions impact user performance. To test performance it is important to select or design a collaborative environment that provides both a high degree of interaction between the user and the computer agent, but also provides performance metrics. To simplify the design of test environments, research could focus on domain specific activities (command and control, personal shopping

assistant, education, etc.). One metric that is important, but often difficult to measure is the level of interaction between the user and computer agent. This is important in order to understand the impact the computer agent has on user performance. A line of research could focus specifically on designing tasks and environments that enhance interactions and simplify data collection.

Titles and Expertise

A possible method for influencing user perception and performance is to give computer agents names, titles or visual clues that identify the expertise of the computer agent. Although some work in this area has been accomplished by Trappl and Petta (1997) there are additional opportunities to determine how a computer agent with identified expertise influences the level of interaction and user performance.

There exist many more opportunities for future research related to designing computer agents with personality and determining their impact on user performance. The goal of this line of research should focus on building an understanding so computer agents can be designed to provide more natural human-computer interaction to improve operator performance.

15. CONCLUSIONS

The results of this study have provided a look into the possibilities of using emotional expressions as a way to incorporate software agents with personality into human-machine collaborative systems. The study provided insight into user perception of personality of static digital faces with emotional expression.

The first phase of the study collected data on the perceived personality of two sets of five digitally generated static faces. The analysis showed several statistically significant relationships exist between participant GENDER, the type of face (FACE SET), the emotional EXPRESSION and the PERSONALITY TRAIT derived from the participant RATINGS. The conclusions we make from these results are that participants not only perceived personality in digitally generated faces, but also perceived different personalities based on the emotional expression of the faces. This knowledge provides an opportunity for system developers and future researchers to use emotion as a way to design and enhance the perceived personality of computer agents.

Before moving into the second phase of research, three faces were selected from the ten that were evaluated in Phase I. The first FACES selected were the two with the EXPRESSION of Joy (A3, B3). These two were selected because their personalities as identified by participant RATINGS most closely matched the personality profile of what previous subjects in a study by Prabhala and Gallimore (2005b,c) indicated they would like to see in an ideal collaborative partner. This personality is identified as Extraverted, Agreeable and Conscientious, with low Neuroticism and moderate Openness. For contrast a third face with a Sad EXPRESSION (A4) was selected because its RATING indicated personality profile opposite that identified for an ideal collaborative partner.

The second phase of this study paired participants with each of the three FACES in order to complete a series of three survival scenarios where the participant was asked to rank a list of items while receiving suggestions from a stimulus FACE. Results for this phase of the study showed that NATIONALITY significantly influenced task SCORE, but did not impact SUGGESTION USE. The analysis showed that American participants scored better than Indian participants. This difference may be attributed to Indian participants being unfamiliar with the activity of ranking items for a survival scenario. In addition, GENDER did not influence SCORE, but did impact SUGGESTION USE. Analysis showed that Female participants were more likely to use the suggestions provided by the computer agent. This result may correspond to gender stereotypes that suggest that women are more open to suggestions.

The FACE stimuli did not have a significant impact on either task SCORE or SUGGESTION USE. This result may indicate that the FACE of a computer agent does not influence performance, but could also be a consequence of no real interaction with the agent during the task used in this study.

The most influential element in participant performance for these survival tasks was the level of SUGGESTION ACCURACY. SUGGESTION ACCURACY influenced both task SCORE and SUGGESTION USE. When participants received Good suggestions the SUGGESTION USE increased and in turn their SCORE improved. This suggests that participants were influenced more by the information they received than by the computer agent delivering the information.

The prospects for advancing collaborative computer agents that provide productive working relationships with human operators are promising. This research has

taken some of the essential first steps for incorporating personality in the design of computer agents using non-verbal facial expression. Although practical application is still a goal for the future, the use of facial expressions as a means to convey personality of a computer agent with additional verbal and nonverbal behaviors appears to be a very real possibility.

16. APPENDICES

APPENDIX A:
BIG-FIVE FACTOR, CENTRAL TRAIT DEFINITIONS

Extraversion:

Extraversion is marked by pronounced engagement with the external world. Extraverts enjoy being with people, are full of energy, and often experience positive emotions. They tend to be enthusiastic, action-oriented, individuals who are likely to say "Yes!" or "Let's go!" to opportunities for excitement. In groups they like to talk, assert themselves, and draw attention to themselves.

Introverts lack the exuberance, energy, and activity levels of extraverts. They tend to be quiet, low-key, deliberate, and disengaged from the social world. Their lack of social involvement should not be interpreted as shyness or depression. The introvert simply needs less stimulation than an extravert does and prefers to be alone. The independence and reserve of the introvert is sometimes mistaken as unfriendliness or arrogance. In reality, an introvert who scores high on the agreeableness dimension will not seek others out but will be quite pleasant when approached.

Agreeableness:

Agreeableness reflects individual differences in concern with cooperation and social harmony. Agreeable individuals value getting along with others. They are therefore considerate, friendly, generous, helpful, and willing to compromise their interests with others'. Agreeable people also have an optimistic view of human nature. They believe people are basically honest, decent, and trustworthy.

Disagreeable individuals place self-interest above getting along with others. They are generally unconcerned with others' well being, and therefore are unlikely to extend themselves for other people. Sometimes their skepticism about others' motives causes them to be suspicious, unfriendly, and uncooperative.

Agreeableness is obviously advantageous for attaining and maintaining popularity. Agreeable people are better liked than disagreeable people. On the other hand, agreeableness is not useful in situations that require tough or absolute objective decisions. Disagreeable people can make excellent scientists, critics, or soldiers.

Conscientiousness:

Conscientiousness concerns the way in which we control, regulate, and direct our impulses. Impulses are not inherently bad. Occasionally time constraints require a snap decision, and acting on our first impulse can be an effective response. Also, in times of play rather than work, acting spontaneously and impulsively can be fun. Impulsive individuals can be seen by others as colorful, fun-to-be-with, and zany.

Nonetheless, acting on impulse can lead to trouble in a number of ways. Some impulses are antisocial. Uncontrolled antisocial acts not only harm other members of society, but also can result in retribution toward the perpetrator of such impulsive acts. Another problem with impulsive acts is that they often produce immediate rewards but undesirable, long-term consequences. Examples include excessive socializing that leads to being fired from one's job, hurling an insult that causes the breakup of an important relationship, or using pleasure-inducing drugs that eventually destroy one's health.

Impulsive behavior, even when not seriously destructive, diminishes a person's effectiveness in significant ways. Acting impulsively disallows contemplating alternative courses of action, some of which would have been wiser than the impulsive choice. Impulsive behavior also sidetracks people during projects that require organized sequences of steps or stages. Accomplishments of an impulsive person are therefore small, scattered, and inconsistent.

A hallmark of intelligence, what potentially separates human beings from earlier life forms, is the ability to think about future consequences before acting on an impulse. Intelligent activity involves contemplation of long-range goals, organizing and planning routes to these goals, and persisting toward one's goals in the face of short-lived impulses to the contrary. The idea that intelligence involves impulse control is nicely captured by the term prudence, an alternative label for the Conscientiousness domain. Prudent means both wise and cautious. In fact, others perceive persons who score high on the Conscientiousness scale as intelligent.

The benefits of high conscientiousness are obvious. Conscientious individuals avoid trouble and achieve high levels of success through purposeful planning and persistence. Others also positively regard them as intelligent and reliable. On the negative side, they can be compulsive perfectionists and workaholics. Furthermore, extremely conscientious individuals might be regarded as stuffy and boring. Unconscientious people may be criticized for their unreliability, lack of ambition, and failure to stay within the lines, but they will experience many short-lived pleasures and they will never be called stuffy.

Neuroticism:

Freud originally used the term neurosis to describe a condition marked by mental distress, emotional suffering, and an inability to cope effectively with the normal demands of life. He suggested that everyone shows some signs of neurosis, but that we differ in our degree of suffering and our specific symptoms of distress. Today neuroticism refers to the tendency to experience negative feelings. Those who score high on Neuroticism may experience primarily one specific negative feeling such as anxiety, anger, or depression, but are likely to experience several of these emotions. People high in neuroticism are emotionally reactive. They respond emotionally to events that would not affect most people, and their reactions tend to be more intense than normal. They are more likely to interpret ordinary situations as threatening, and minor frustrations as hopelessly difficult. Their negative emotional reactions tend to persist for unusually long periods of time, which means they are often in a bad mood. These problems in emotional regulation can diminish a neurotic's ability to think clearly, make decisions, and cope effectively with stress.

At the other end of the scale, individuals who score low in neuroticism are less easily upset and are less emotionally reactive. They tend to be calm, emotionally stable, and free from persistent negative feelings. Freedom from negative feelings does not mean that low scorers experience a lot of positive feelings; frequency of positive emotions is a component of the Extraversion domain.

Openness to Experience:

Openness to Experience describes a dimension of cognitive style that distinguishes imaginative, creative people from down-to-earth, conventional people. Open people are intellectually curious, appreciative of art, and sensitive to beauty. They tend to be, compared to closed people, more aware of their feelings. They tend to think and act in individualistic and nonconforming ways. Intellectuals typically score high on Openness to Experience; consequently, this factor has also been called culture or intellect. Nonetheless, intellect is probably best regarded as one aspect of openness to experience. Scores on openness to experience are only modestly related to years of education and scores on standard intelligent tests.

Another characteristic of the open cognitive style is a facility for thinking in symbols and abstractions far removed from concrete experience. Depending on the individual's specific intellectual abilities, this symbolic cognition may take the form of mathematical, logical, or geometric thinking, artistic and metaphorical use of language, music composition or performance, or one of the many visual or performing arts. People with low scores on openness to experience tend to have narrow, common interests. They prefer the plain, straightforward, and obvious over the complex, ambiguous, and subtle. They may regard the arts and sciences with suspicion, regarding these endeavors as abstruse or of no practical use. Closed people prefer familiarity to novelty; they are conservative and resistant to change.

Openness is often presented as healthier or more mature by psychologists, who are often themselves open to experience. However, open and closed styles of thinking are useful in different environments. The intellectual style of the open person may serve a

professor well, but research has shown that closed thinking is related to superior job performance in police work, sales, and a number of service occupations.

APPENDIX B:
BIG-FIVE FACTOR, PERSONALITY SUB-TRAIT DEFINITIONS

Extraversion:

- Friendliness: Genuinely like other people and openly demonstrate positive feelings toward others.
- Gregariousness: Find the company of others pleasantly stimulating and rewarding; and enjoy the excitement of crowds.
- Assertiveness: Like to speak out, take charge, and direct the activities of others. Tend to be leaders of groups.
- Activity Level: Lead fast-paced, busy lives; move about quickly, energetically, vigorously, and are involved in many activities.
- Excitement-Seeking: Bored without high levels of stimulation. They are likely to take risks and seek thrills.
- Cheerfulness: Typically experience a range of positive feelings, including happiness, enthusiasm, optimism, and joy.

Agreeableness:

- Trust: Assumes that most people are fair, honest, and have good intentions.
- Morality: See no need for pretense or manipulation and are therefore candid, frank, and sincere.
- Altruism: Find helping other people genuinely rewarding and are generally willing to assist those who are in need.

- Cooperation: Dislike confrontations. They are perfectly willing to compromise or to deny their own needs in order to get along with others.
- Modesty: Do not like to claim that they are better than other people.
- Sympathy: Tenderhearted and compassionate. They feel the pain of others vicariously and are easily moved to pity.

Conscientiousness:

- Self-Efficacy: Confidence in one's ability to accomplish things.
- Orderliness: Well organized and like to live according to routines and schedules.
- Dutifulness: Strong sense of duty and moral obligation.
- Achievement-Striving: Strive hard to achieve excellence; drive to be recognized as successful keeps them on track toward lofty goals.
- Self-Discipline: The ability to persist at difficult or unpleasant tasks until they are completed.
- Cautiousness: The disposition to think through possibilities before acting.

Neuroticism:

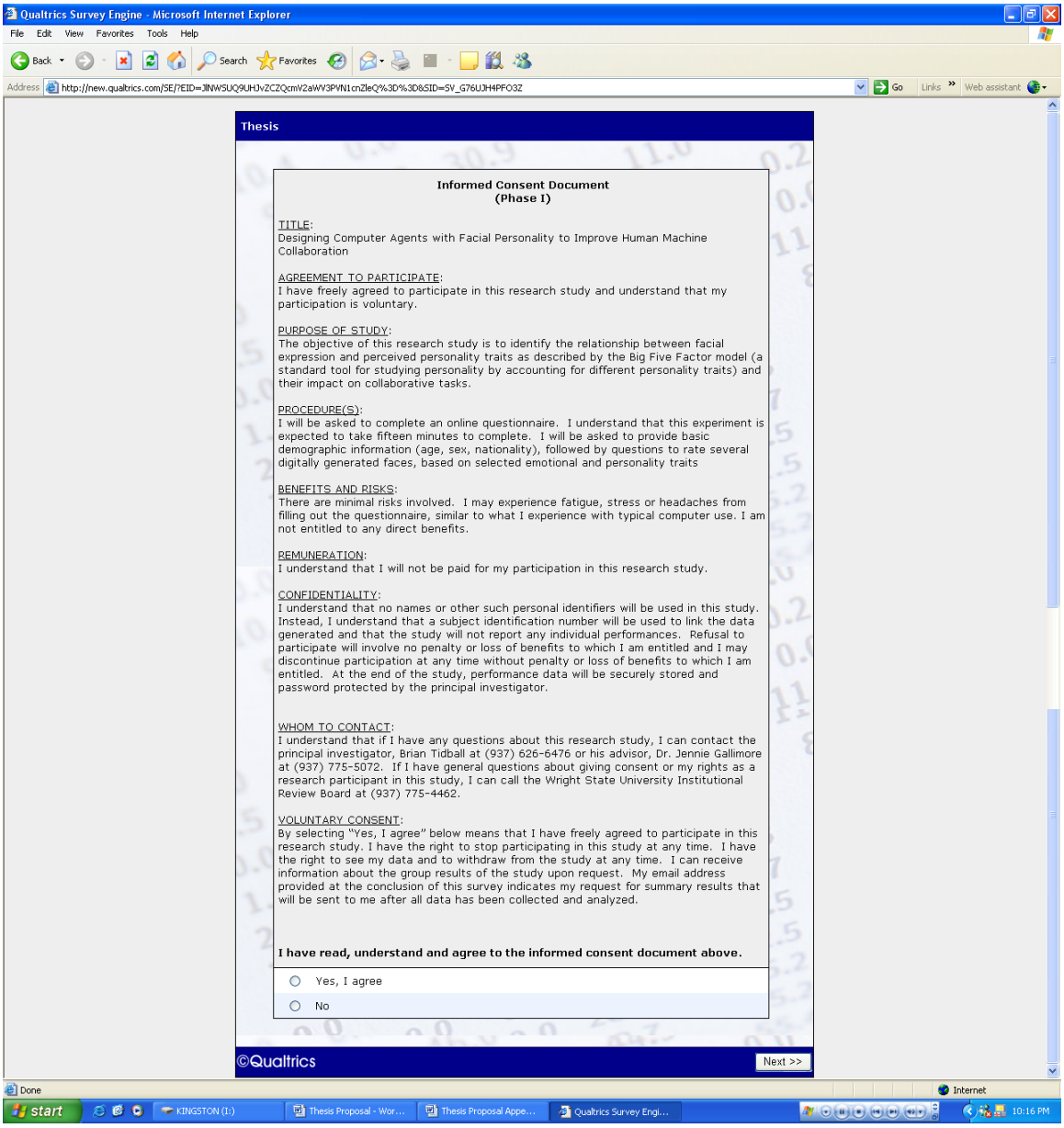
- Anxiety: Often feel like something dangerous is about to happen. They may be afraid of specific situations or be just generally fearful. They feel tense, jittery and nervous.
- Anger: Feel enraged when things do not go their way. They are sensitive about being treated fairly and feel resentful and bitter when they feel they are being cheated.
- Depression: Tendency to feel sad, dejected, discouraged and lack energy.

- Self-Consciousness: Sensitive about what others think of them, feel shy and uncomfortable around others.
- Immoderation: Feel strong cravings and urges that they have difficulty resisting.
- Vulnerability: Experience panic, confusion, and helplessness when under pressure or stress.

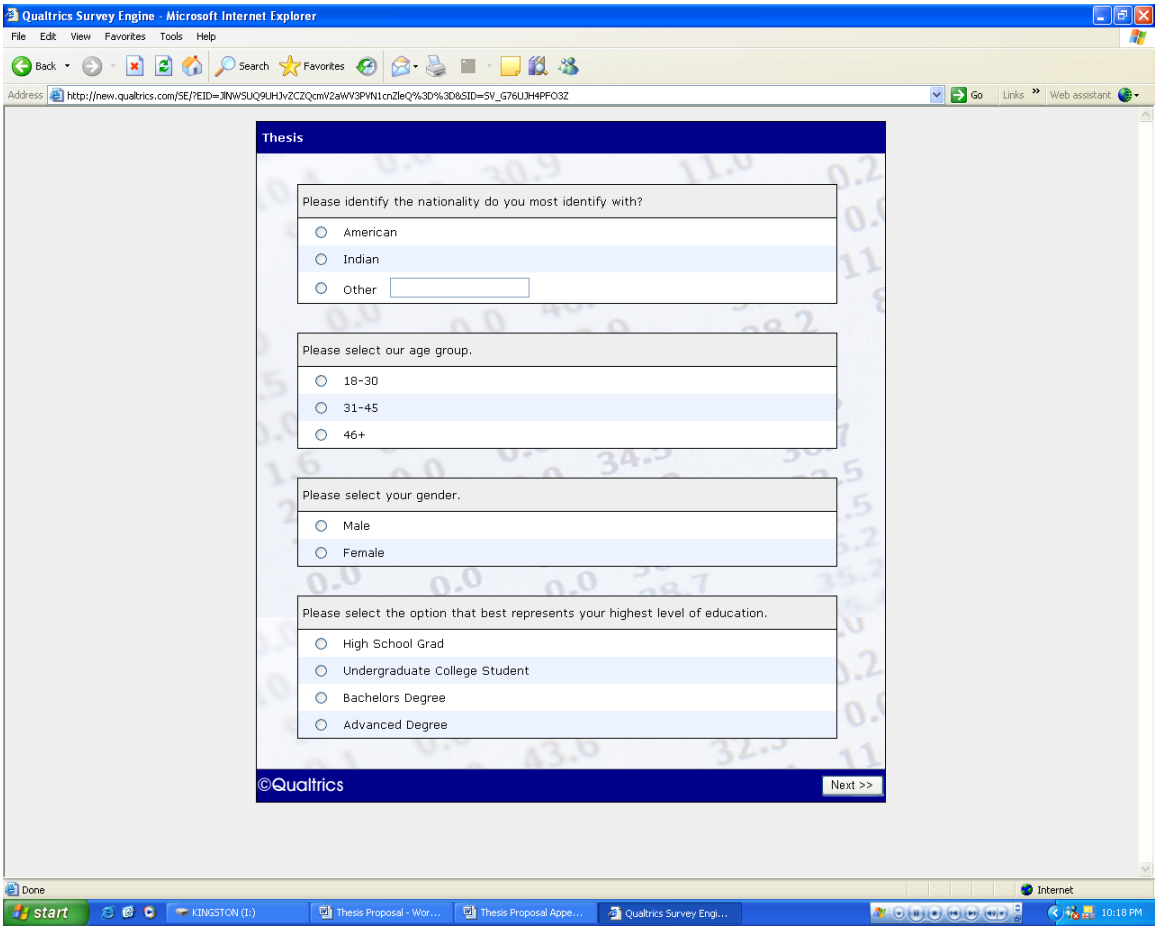
Openness:

- Imagination: The real world is often too plain and ordinary; use fantasy as a way of creating a richer, more interesting world.
- Artistic Interests: Love beauty, both in art and in nature. They become easily involved and absorbed in artistic and natural events.
- Emotionality: Good access to and awareness of their own feelings.
- Adventurousness: Eager to try new activities, travel, and experience different things.
- Intellect: Love to play with ideas, debate intellectual issues and enjoy riddles, puzzles, and brainteasers.
- Liberalism: readiness to challenge authority, convention, and traditional values.

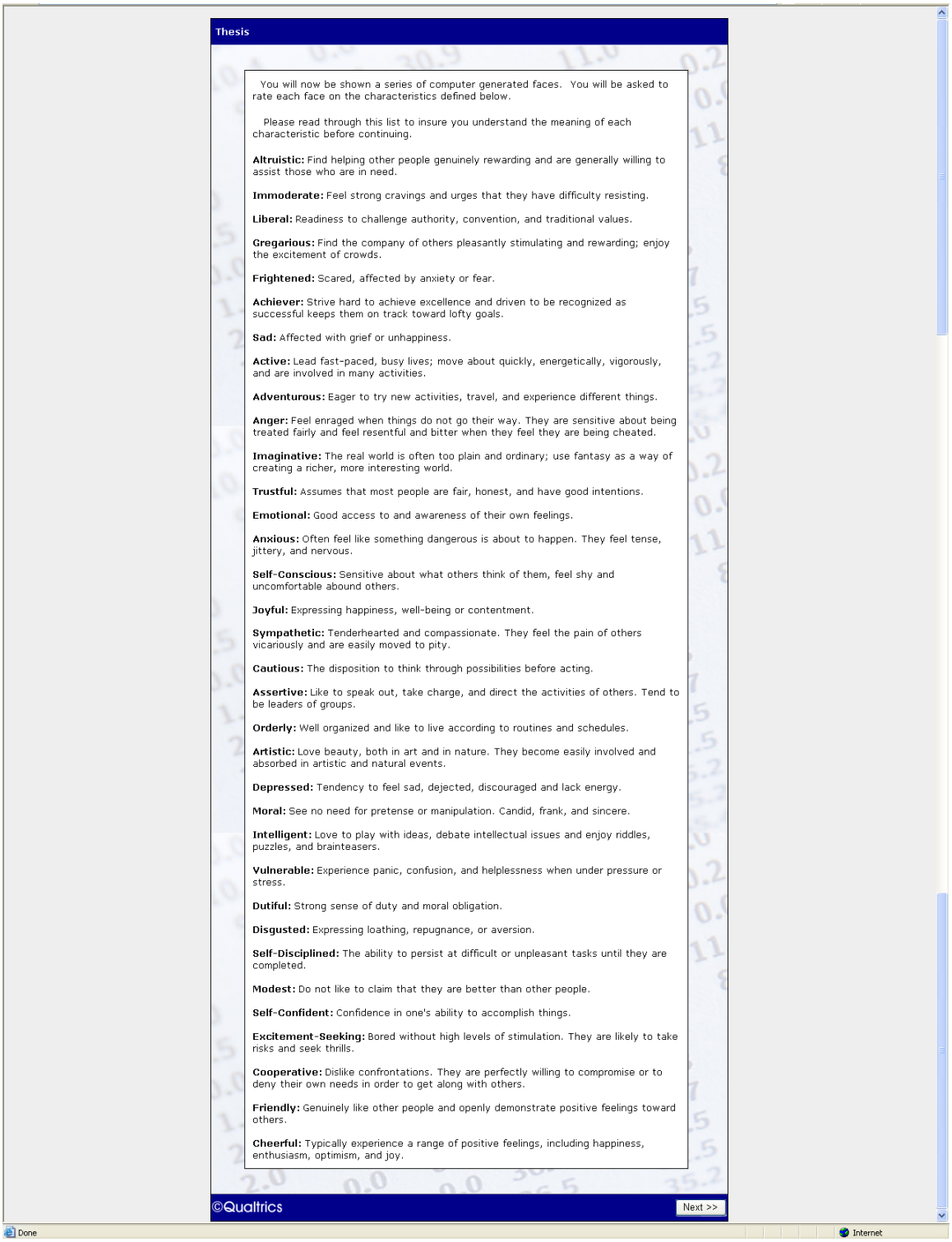
APPENDIX C
PHASE I: CONSENT AGREEMENT



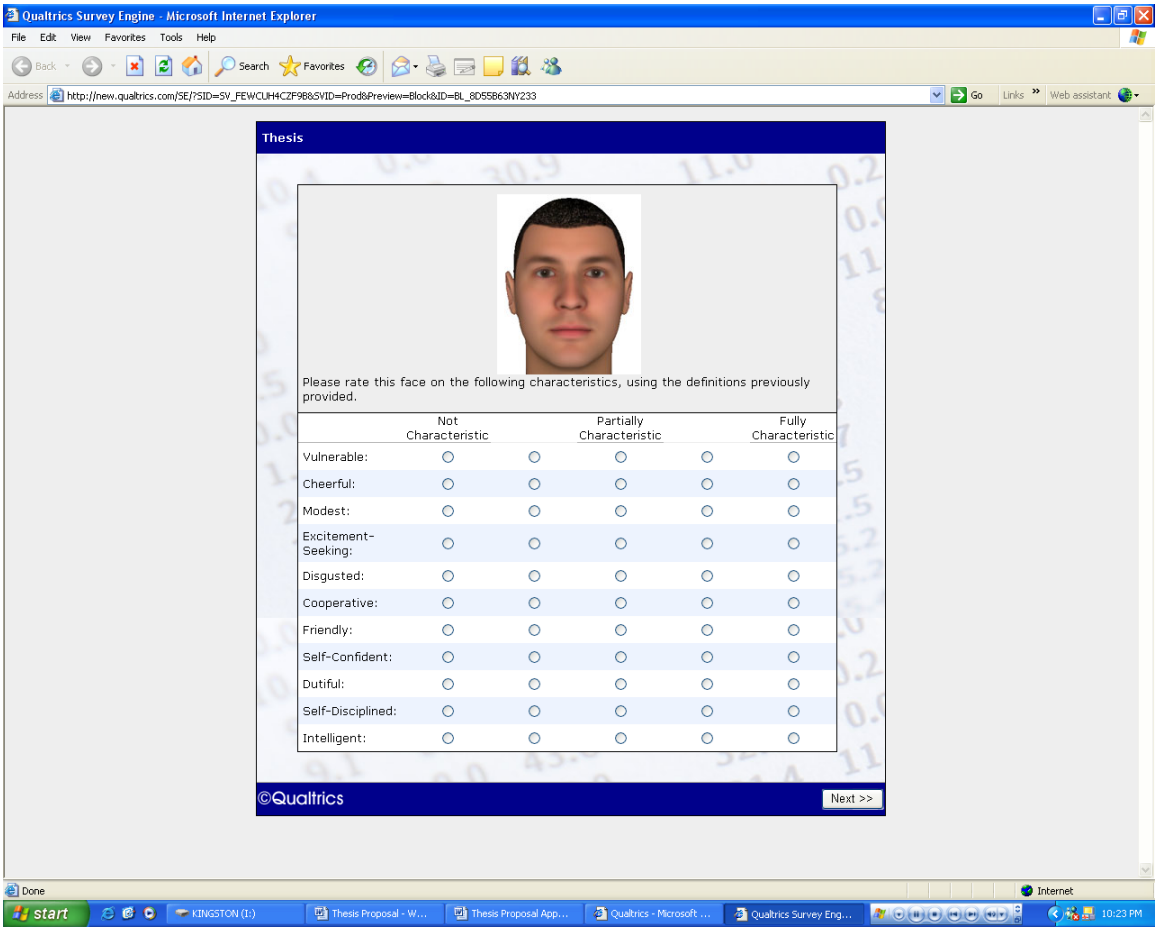
APPENDIX C
PHASE I: DEMOGRAPHIC QUESTIONS



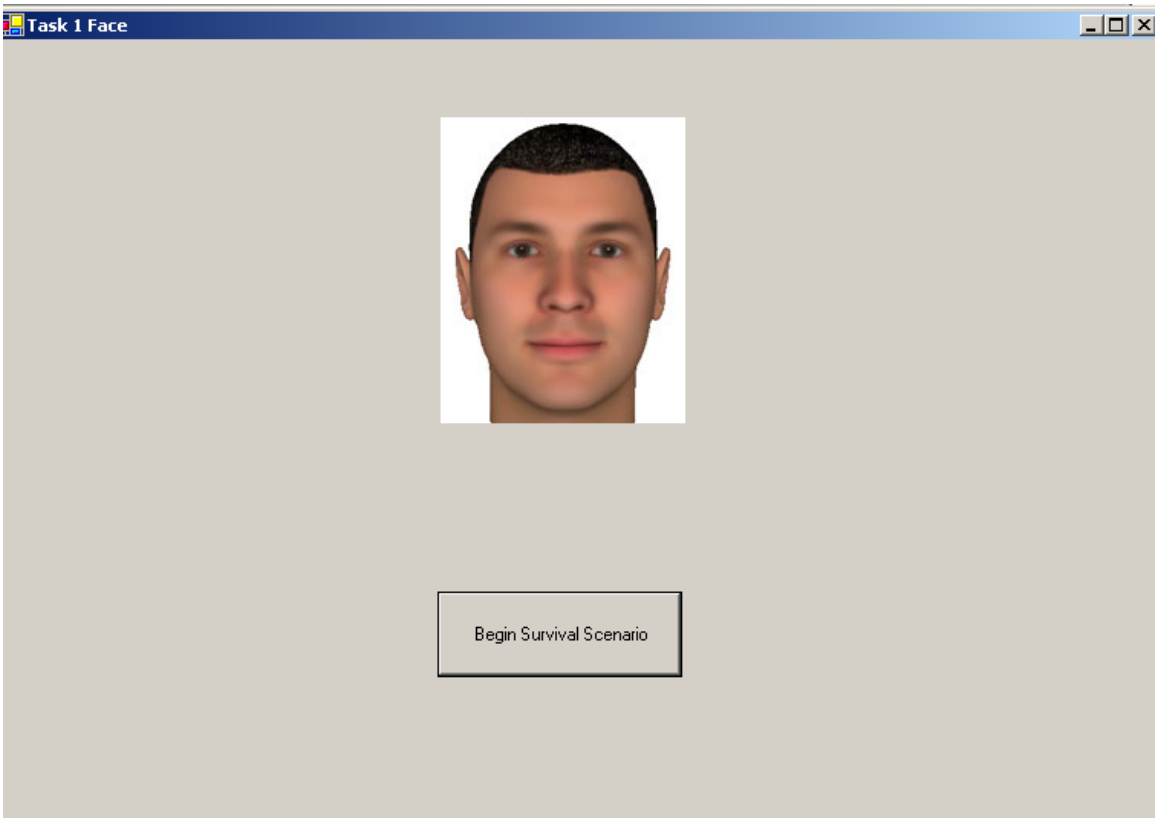
APPENDIX C:
PHASE I: TRAIT DEFINITIONS



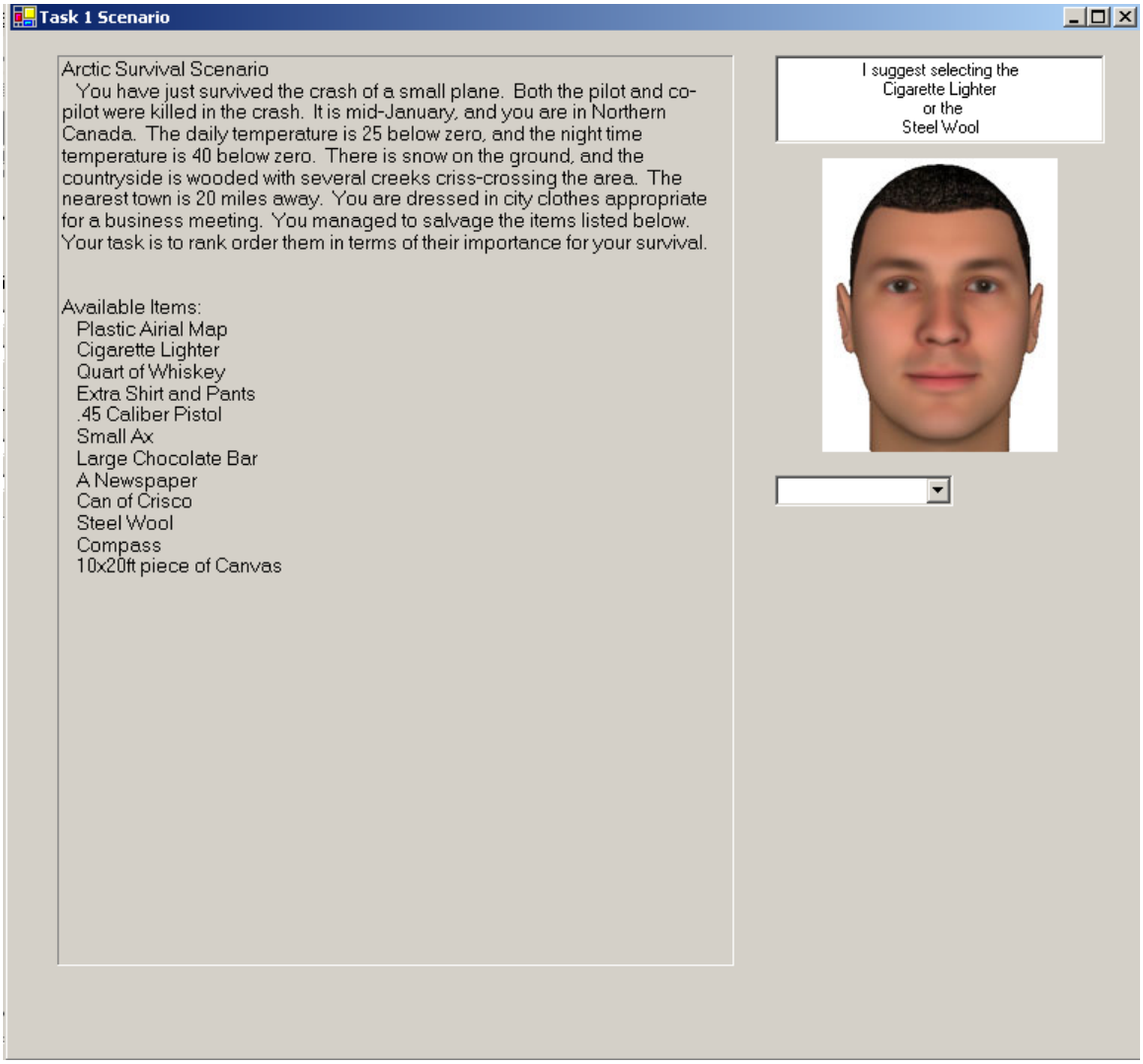
APPENDIX C
PHASE I: SAMPLE TRAIT QUESTIONNAIRE



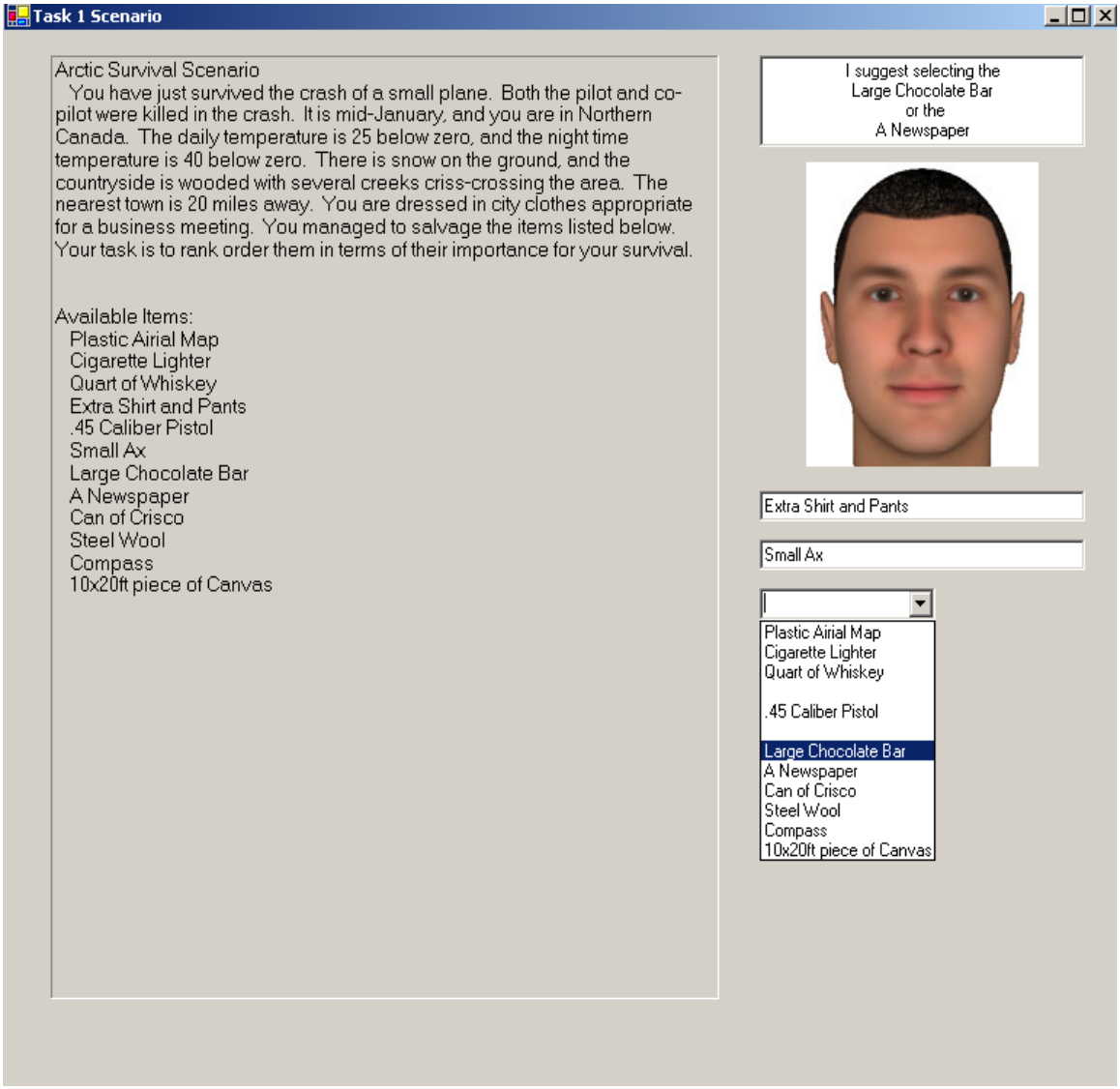
APPENDIX D
PHASE II: INTRODUCTION SCREEN



APPENDIX D:
PHASE II: SURVIVAL SCENARIO SCREEN



APPENDIX D:
PHASE II: SURVIVAL SCENARIO SCREEN IN PROGRESS



APPENDIX E
SHIP WRECKED SCENARIO

Description

After getting caught in a storm in the north Atlantic, you find yourself sheltering in a cove which is about 400 meters from the rocks where the boat was lost. The water bordering the island is very deep (500m); the cliffs and rocks around the edge of the island are bare and dangerous. They range in height from 200 to 300 meters. The beach is about 30 meters long and 8 meters deep, rising sharply into rocky and rugged terrain. The only inhabitants of this very remote island are the huge gulls that nest on the top of some of the higher cliffs. Only occasionally does a cruise or any other boat visit this end of the fjord, which is uninhabited and somewhat foreboding. Air temperature in the area for June is typically 15-22˚C (60-70˚F) during the day and 7-10˚C (45-50˚F) at night, with 12 to 18 knot winds. The relative humidity is about 60-65%. During storms, the temperature is known to drop suddenly some 20 degrees or more. In the very worst summer weather, snow is not unknown!! Water temperature in June is 10 to 13˚C (50-55˚F). Fog is pretty constant in this part of the fjord, and this time of the year brings heavy rainfall and brief storms, which may last up to 48 hours. Sunrise is 3.10 a.m. and sunset is 23.15 p.m. During the wreck you managed to salvage the items listed below. Your task is to rank them in terms of their importance for your crew in allowing them to reach the rendezvous point. Place the number **1** by the most important item, the number **2** by the second most important, and so on through number **12** for the least important.

Ranked List of Available Items

1. **6 x 14 blue canvas canopy** -Shelter and water collection. Parts may be cut from the main canopy to make solar water stills. This and the knife below are the most useful items on this list. With these two items, the chances of survival and rescue reach 86%.
2. **Swiss Army knife with multiple tools** - Best all-round tool for survival. The shiny surface of the blade can act as a signal device.
3. **One dinghy paddle** - Acts as a support for the canvas
4. **Seven cans of food** - Can be eaten, but once opened can also act as water containers.
5. **80' of nylon rope** - Can be used to tie the canvas together - also acts as wicks for lights.
6. **Two orange life jackets** - When set on fire, these will make a large smoke trail which can be seen for many kilometers.
7. **Three flotation cushions** - As with the life jackets, these can be burnt to raise a smoke trail.
8. **One scuba mask** - The glass can be used as a light signal by reflecting the sun's rays. Might also be useful to smash and have additional knives. The water is so deep and cold death by hypothermia would be inevitable for anyone attempting to use the mask to swim.
9. **Various pieces of rigging wire** - Might try to use these for traps, but frankly, the chances of catching anything large enough to reward the effort is unlikely.

10. **One standard-sized plastic cooler** - Good for storing water and fresh food - but since both will be very rare, you might use it as a basket to carry firewood from the trees and driftwood.
11. **One pair of swim fins -**Could be used for digging or as smoke-producing fuel. The water is so deep and cold that death by hypothermia would be inevitable for anyone attempting to use the fins to swim.
12. **Four bottles of suntan lotion** - You could use these as fuel to raise smoke from the fire or as fuel for wick-based lights, using lengths of the nylon cord as a wick.

APPENDIX F
MOON SURVIVAL SCENARIO

Description

You are a member of a space crew originally scheduled to rendezvous with a mother ship on the lighted surface of the moon. However, due to mechanical difficulties, your ship was forced to land at a spot some 200 miles from the rendezvous point. During reentry and landing, much of the equipment aboard was damaged and, since survival depends on reaching the mother ship, the most critical items available must be chosen for the 200-mile trip. Below are listed the 15 items left intact and undamaged after landing. Your task is to rank order them in terms of their importance for your crew in allowing them to reach the rendezvous point. Place the number **1** by the most important item, the number **2** by the second most important, and so on through number **12** for the least important.

Ranked List of Available Items

1. **Two 100 lb. tanks of oxygen** - Most pressing survival need (weight is not a factor since gravity is one-sixth of the Earth's -- each tank would weigh only about 17 lbs. on the moon)
2. **5 gallons of water** - Needed for replacement of tremendous liquid loss on the light side
3. **Stellar map** - Primary means of navigation - star patterns appear essentially identical on the moon as on Earth

4. **Food concentrate** - Efficient means of supplying energy requirements
5. **Solar-powered FM receiver-transmitter** - For communication with mother ship (but FM requires line-of-sight transmission and can only be used over short ranges)
6. **50 feet of nylon rope** - Useful in scaling cliffs and tying injured together
7. **First aid kit, including injection needle** - Needles connected to vials of vitamins, medicines, etc. will fit special aperture in NASA space suit
8. **Parachute silk** - Protection from the sun's rays
9. **Self-inflating life raft** - CO_2 bottle in military raft may be used for propulsion
10. **Signal flares** - Use as distress signal when the mother ship is sighted
11. **Two .45 caliber pistols** - Possible means of self-propulsion
12. **One case of dehydrated milk** - Bulkier duplication of food concentrate

APPENDIX G
ARCTIC SURVIVAL SCENARIO

Description

You have just survived the crash of a small plane. Both the pilot and co-pilot were killed in the crash. It is mid-January, and you are in Northern Canada. The daily temperature is 25 below zero, and the night time temperature is 40 below zero. There is snow on the ground, and the countryside is wooded with several creeks criss-crossing the area. The nearest town is 20 miles away. You are dressed in city clothes appropriate for a business meeting. You managed to salvage the items listed below. Your task is to rank them in terms of their importance for your survival. Place the number **1** by the most important item, the number **2** by the second most important, and so on through number **12** for the least important.

Ranked List of Available Items

1. **Cigarette lighter** (without fluid) - The gravest danger facing the group is exposure to cold. The greatest need is for a source of warmth and the second greatest need is for signaling devices. This makes building a fire the first order of business. Without matches, something is needed to produce sparks, and even without fluid, a cigarette lighter can do that.
2. **Ball of steel wool** - To make a fire, the survivors need a means of catching the sparks made by the cigarette lighter. This is the best substance for catching a spark and supporting a flame, even if the steel wool is a little wet.

3. **Extra shirt and pants for each** - Besides adding warmth to the body, clothes can also be used for shelter, signaling, bedding, bandages, string (when unraveled), and fuel for the fire.
4. **Can of Crisco shortening** - This has many uses. A mirror-like signaling device can be made from the lid. After shining the lid with steel wool, it will reflect sunlight and generate 5 to 7 million candlepower. This is bright enough to be seen beyond the horizon. While this could be limited somewhat by the trees, a member of the group could climb a tree and use the mirrored lid to signal search planes. If they had no other means of signaling than this, they would have a better than 80% chance of being rescued within the first day. There are other uses for this item. It can be rubbed on exposed skin for protection against the cold. When melted into oil, the shortening is helpful as fuel. When soaked into a piece of cloth, melted shortening will act like a candle. The empty can is useful in melting snow for drinking water. It is much safer to drink warmed water than to eat snow, since warm water will help retain body heat. Water is important because dehydration will affect decision-making. The can is also useful as a cup.
5. **20 x 20 foot piece of canvas** - The cold makes shelter necessary, and canvas would protect against wind and snow (canvas is used in making tents). Spread on a frame made of trees, it could be used as a tent or a wind screen. It might also be used as a ground cover to keep the survivors dry. Its shape, when contrasted with the surrounding terrain, makes it a signaling device.
6. **Small ax** - Survivors need a constant supply of wood in order to maintain the fire. The ax could be used for this as well as for clearing a sheltered campsite,

cutting tree branches for ground insulation, and constructing a frame for the canvas tent.

7. **Family size chocolate bars** (one each) - Chocolate will provide some food energy. Since it contains mostly carbohydrates, it supplies the energy without making digestive demands on the body.
8. **Newspapers** (one each) - These are useful in starting a fire. They can also be used as insulation under clothing when rolled up and placed around a person's arms and legs. A newspaper can also be used as a verbal signaling device when rolled up in a megaphone-shape. It could also provide reading material for recreation.
9. **Loaded .45-caliber pistol** - The pistol provides a sound-signaling device. (The international distress signal is three shots fired in rapid succession). There have been numerous cases of survivors going undetected because they were too weak to make a loud enough noise to attract attention. The butt of the pistol could be used as a hammer, and the powder from the shells will assist in fire building. By placing a small bit of cloth in a cartridge emptied of its bullet, one can start a fire by firing the gun at dry wood on the ground. The pistol also has some serious disadvantages. Anger, frustration, impatience, irritability, and lapses of rationality may increase as the group awaits rescue. The availability of a lethal weapon is a danger to the group under these conditions. Although a pistol could be used in hunting, it would take an expert marksman to kill an animal with it. Then the animal would have to be transported to the crash site, which could prove difficult to impossible depending on its size.

10. **Quart of 100 proof whiskey** - The only uses of whiskey are as an aid in fire building and as a fuel for a torch (made by soaking a piece of clothing in the whiskey and attaching it to a tree branch). The empty bottle could be used for storing water. The danger of whiskey is that someone might drink it, thinking it would bring warmth. Alcohol takes on the temperature it is exposed to, and a drink of minus 30 degrees Fahrenheit whiskey would freeze a person's esophagus and stomach. Alcohol also dilates the blood vessels in the skin, resulting in chilled blood belong carried back to the heart, resulting in a rapid loss of body heat. Thus, a drunken person is more likely to get hypothermia than a sober person.
11. **Compass** - Because a compass might encourage someone to try to walk to the nearest town, it is a dangerous item. Its only redeeming feature is that it could be used as a reflector of sunlight (due to its glass top).
12. **Sectional air map made of plastic** - This is also among the least desirable of the items because it will encourage individuals to try to walk to the nearest town. Its only useful feature is as a ground cover to keep someone dry.

17. REFERENCES

Bahr, G.S., Wheeler-Atkinson, B.F., Walwanis-Nelson, M.M. and Stewart E. (2005). Augmented cognition and time perception. In *Proceedings of the 11th International Conference in Human Computer Interaction*, *AUGCOG Conference* July 22-27, Las Vegas, NV.

Baker, D. P., and Salas, E. (1992). Principles for measuring teamwork skills. *Human Factors*, *34*(4), pp. 469-475.

Berry D.S. (1989). Were the physiognomists right? Personality correlates of facial babyshyness. *Personality and Social Psychology Bulletin*,*15*(2), 266-279.

Borgatta, E.F. (1964). The structure of personality characteristics. *Behavioral Science, 9*, 8-17.

Buss, D. M. (1996). Social adaptation and five major factors of personality. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. 180-207). New York: Guilford.

Cannon-Bowers, J.A. and Salas, E. (1998). Individual and team decision making under stress: Theoretical underpinnings. In J.A.Cannon-Bowers and E. Salas (Eds.) *Making decisions under stress: Implication for individual and team training.* (pp. 17-38).Washington, D.C., American Psychological Association Press.

Charlesworth, W.R. and Kreutzer, M.A. (1973). Facial expression of infants and children. In P. Ekman, Ed., *Darwin and Facial Expression.* New York: Academic Press. pp. 91-168.

Digman, J.M. (1997). Higher-order factors of the big five. *Journal of Personality and Social Psychology*, *73*(6), 1246-1256.

Digman, J.M. and Takemoto-Chock, N.K. (1981). Factors in the natural language of personality: Re-analysis, comparison and interpretation of six major studies. *Multivariate Behavioral Research*, *16*, 149-170.

Ekman, P. (1979). About brows: Emotional and conversational signals. In M. Cranach, K. Foppa, W. Lepenies, and D. Ploog (Eds.). *Human Ethology* claims and limits of a new discipline: Contributions to the colloquium. Cambridge, UK: von Cambridge University Press. pp.169–202.

Ekman, P. and Friesen, W.V. (1976). Measuring facial movement. *Environmental Psychology and Nonverbal Behavior*, *1*, 56-75.

Ekman, P., Sorenson, E.R. and Frieson, W.V. (1969). Pan-cultural elements in facial emotion. *Science*, *164*, 86-88.

Essa, I. and Pentland, A (1997). Coding, analysis, interpretation, and recognition of facial expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 757-763.

Eitelman, S.M., Wheeler-Atkinson, B.F., Walwanis-Nelson, M.M. and Stiso, M.E. (2005). Enhancing the common distributed mission training station through augmented cognition technology. In *Proceedings of the 11th International Conference in Human Computer Interaction*, *AUGCOG Conference* July 22-27, Las Vegas, NV.

Fogg, B.J., and Tseng, H. (1999). The elements of computer credibility. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. pp. 80-87.

Galton, F. (1884). Measurement of character. *Fortnightly Review. 36*, pp. 179-185.

Goldberg, L. R. (1990). An alternative "description of personality": the big five factor structure. *Journal of Personality and Social Psychology*, *59*(6), 1216-1229.

Goldberg, L.R. (1999). A broad-bandwidth public domain, personality inventory measuring the lower-level facets of several five-factor models. In L. Mervielde, L. Deary, F. De Fruyt, and F. Ostendorf (Eds.), *Personality Psychology in Europe* Vol. 9, (pp. 7-28). Tilburg, The Netherlands: Tilburg University Press.

Goren, D and Wilson, H.R. (2006). Quantifying facial expression recognition across Viewing Conditions.*Vision Research*, 46(8-9),1253-62.

Hagar, J.C. and Ekman, P. (1983). The inner and outer meanings of facial expressions. *Social Psychology: A Sourcebook.* New York: The Guilford Press

Hassin, R. and Trope, Y. (2000). Facing faces: studies on the cognitive aspects of physiognomy. *Journal of Personality and Social Psychology, 78*(5), 837-852.

Hoc, J. (2001). Towards a cognitive approach to human-machine cooperation in dynamic situations. *International Journal of Human-Computer Studies*, *54*, 509-540.

http://scoutingweb.com/scoutingweb/SubPages/SurvivalGame.htm

http://www.bpccs.com/lcas/Articles/survive.htm

http://www.ideels.uni-bremen.de/narg_situation.html

http://www.m-w.com/

Izard, C.E. (1971). *The face of emotion.* New York: Appleton-Century-Crofts.

Johnson, M., Kulkarni, K., Raj, A.,Carff, R. and Bradshaw, J.M. (2005). AMI: An adaptive multi-agent framework for augmented cognition. In *Proceedings of the 11th International Conference in Human Computer Interaction*, *AUGCOG Conference* July 22-27, Las Vegas, NV.

Kent, Jeffrey (2002). *Visual Basic .NET: A Beginner's Guide.* Berkley, CA: McGraw-Hill/Osborne.

Liu, Z. (2003). A fully automatic system to model faces from a single image. Microsoft Research, Technical Report MSR-TR-2003-55.

Liu, Z., Shan, Y. Zhang, Z.(2001). Expressive expression mapping with ratio images. In *SIGGRAPH 2001*, Los Angeles, August 12-17, pp 271-276.

Markopoulos, P., Ruyter, B., Privender, S. and Breemen, A. (2005). Case study bringing social intelligence into home dialog systems. *Interactions*,*12*(4), 37-44.

Matsumoto, D. (1987). The role of facial response in the experience of emotion: More methodological problems and a meta-analysis. *Journal of Personality and Social Psychology, 52*(4), 769-774.

McCrae, R.R., and Costa, P.T., Jr. (1985). Comparison of EPI and psychoticism scales with measures of the five-factor model of personality. *Personality and Individual Differences, 6*,587-597.

McCrae, R.R., and Costa, P.T., Jr. (1997). Personality trait structure as a human universal. *American Psychologist*, *52*(5), 509-516.

McCrae, R.R., and Costa, P.T., Jr. (1999). A five-factor theory of personality. In L.A. Pervin and O.P. John (Eds.), *Handbook of personality: theory and research* (pp. 139-153). NY: The Guilford Press.

Millot, P., and Lemoine, M.P. (1998). An attempt for generic concepts toward human-machine cooperation. *IEEE International Conference on Systems, Man, and Cybernetics*, *1*,1044-1049.

Muir, B.M. (1987). Trust between humans and machines, and the design of decision aids. *International Journal of Man-Machine Studies*, *27*, 527-539.

Nass, C., Fogg, B.J., and Moon, Y. (1996). Can computers be teammates? *International Journal of Human-computer Studies*, *45*, 669-678.

Nass, C., Steuer, J., Tauber, E., and Reeder, D. C. (1993). Anthropomorphism, agency and ethopoeia: Computers as social actors. In *Proceedings of the International CHI Conference*, Amsterdam, The Netherlands, pp. 111-112.

Prabhala, S., Gallimore, J. (2005a). Perceptions of personality in computer agents: Effects of culture and gender. In *Proceedings of the 2005 Human Factors and Ergonomics 49th Annual Meeting,* Orlando, FL, September 26-30, pp 716-720

Prabhala, S., Gallimore, J. (2005b). Can humans perceive personality in computer agents? In *Proceedings of the Institute of Industrial Engineers Annual Conference*, May 14-18, Atlanta, paper 317.

Prabhala, S., Gallimore, J. (2005c). Developing computer agents with personality. In *Proceedings of the 11th International Conference in Human Computer Interaction*, *AUGCOG Conference* July 22-27, Las Vegas, NV.

Pew, R.W., & Mavor, A.S. (1998). *Modeling human and organizational behavior: Applications to military simulation*. Washington, D.C.: National Academy Press.

Reeves, B., and Nass, C. (1996). The media equation: How people treat computers, Television and new media like real people and places. New York: Cambridge University Press.

Ruckmick, C.A. (1921). A preliminary study of the emotions, *Psychology Monograghs*, *30*, 30-36.

Saucier, G., and Goldberg, L.R. (1998). What is beyond the big five? *Journal of Personality*, *66*(4), 495-524.

Schmorrow, D.D., and Kruse, A.A. (2002). Improving human performance through advanced cognitive system technology. In *Proceedings of the 24th Interservice/ Industry Training Simulation and Education Conference,* Orlando, FL.

Silverman, B. G., Cornwell, J., and O'Brien, K. (2003). Human performance simulation. In Ness J.W., Ritzer, J.W. and Tepe V. (Eds.), *Metrics and methods in human performance research toward individual and small unit simulation.* Washington, D.C.: Human Systems Information Analysis Center.

Soldz, S., and Vaillant, G.E. (1999). The big five personality traits and the life course: A 45 year longitudinal study. *Journal of Research in Personality*. *33*, 208-232.

Stall, J.,Creighton, L. and Lehman, A. (2005). *JMP Start Statistics: A guide to statistics and data analysis using JMP and JMP IN Software.* SAS Institute Inc.

Thurstone, L.L. (1934). The vectors of mind. *Psychological Review*,*41*,1-32.

Trappl, R., and Petta, P., (Eds.), (1997). Creating personalities for synthetic actors: Towards autonomous personality agents. New York: Springer.

Watson, J.M. (2004). From interpretation to identification; a history of facial images in the sciences of emotion. *History of the human sciences*, *17*(1), 29-51.

Wilson, H. R., Loffler, G. and Wilkinson, F. (2002). Synthetic faces, face cubes, and the geometry of face space. *Vision Research, 42*, 2909-2923.

Winter, D.G., and Barenbaum, N.B. (1999). History of modern personality theory and research. In L.A. Pervin and O. P. John (Eds.), *Handbook of personality: Theory and research* (pp. 3-27). NY: Guilford Press.

Zebrowitz, L.A. (1997.) *Reading faces: Window to the soul?* Boulder, CO: Westview Press.