Gender and Personality Trait Measures Impact Degree of Affect Change in a Hedonic Computing Paradigm

Jeremy D. Schwark a, Igor Dolgov a, Daniel Hor a & William Graves a

a Psychology Department, New Mexico State University, Las Cruces, New Mexico, USA


To cite this article: Jeremy D. Schwark, Igor Dolgov, Daniel Hor & William Graves (2013): Gender and Personality Trait Measures Impact Degree of Affect Change in a Hedonic Computing Paradigm, International Journal of Human-Computer Interaction, 29:5, 327-337

To link to this article: http://dx.doi.org/10.1080/10447318.2012.711703

Full terms and conditions of use: http://www.tandfonline.com/page/terms-and-conditions

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
Gender and Personality Trait Measures Impact Degree of Affect Change in a Hedonic Computing Paradigm

Jeremy D. Schwark, Igor Dolgov, Daniel Hor, and William Graves
Psychology Department, New Mexico State University, Las Cruces, New Mexico, USA

To date, affective computing research has acknowledged individual differences with regard to detecting affect, yet little research has explored how these individual differences may determine the degree to which affective computing is successful in manipulating the affect of specific computer users. The current study used individual difference measures to predict how much an individual can be influenced by a hedonic computing paradigm: a simple trivia game. Female participants responded in a greater way to positive affective feedback about their performance than did men. Moreover, several personality traits, including neuroticism, narcissism, self-esteem, and extraversion, augmented the degree to which affect changed as a result of playing the game. The results are consistent with the gender differences hypothesis, and the authors conclude that individual differences, particularly gender and personality traits, play a large role in the potential impact of computing platforms and would be useful in personalizing the affective nature of the human–computer interaction.

Affective computing, the capability of a computer to sense a user's emotional state and respond similarly, has become a major topic of human–computer interaction (HCI) and human factors research. Because computers are a cornerstone in so many aspects of modern life, affective computing methods have value in a wide range of domains, such as education and training, quality of service, synthetic avatars and intelligent agents, graphical user interfaces, virtual environments, and smart house technologies, among others (Beale & Creed, 2009; McNeese, 2003). Successfully implementing affective computing in a platform is not without its challenges (Carberry & de Rosis, 2008; Picard, 1997, 2003), the greatest of which is programming a machine to have the ability to successfully sense human emotion. Although difficult, the goal is not insurmountable, and software architectures that adapt to such needs and concerns have already been created (Burleson, Picard, Perlin, & Lippincott, 2004; Neviarouskaya, Prendinger, & Ishizuka, 2010; Picard, 1997, 2003). Although progress is being made in these research and design areas, little attention has been paid to the subjective impact of affective computing manipulations on the users.

To date, most research in affective computing has focused on detecting and expressing emotion (Picard, 2003), although an emphasis on understanding, interpreting, and experiencing emotion has been expressed as well (Boehner, DePaula, Dourish, & Sengers, 2007). In line with this reasoning, accounting for individual differences in affective response among the user population needs to be empirically addressed. For example, it is acknowledged that people respond differentially to stressful stimuli. To detect stress and affect, methods that cater to the individual have been used, such as measuring galvanic skin response (Healey, Picard, & Dabek, 1998), blood volume pressure (Ward & Marsden, 2003), electromyograms (Scheirer, Fernandez, Klein, & Picard, 2002), squeeze pressure on a mouse (van Galen, Liesker, & De Haan, 2004), keystrokes (Vizer, Zhou, & Sears, 2009), and response time (Lemmens, De Haan, van Galen, & Meulenbroek, 2007), among others. These physiological measurements are customized to the user and have even been used to alter computer game difficulty in real time based on the user’s current affect (Liu, Agrawal, Sarkar, & Chen, 2009). Unfortunately, this acknowledgment of individual differences only stretches as far as detecting affect, and a “one size fits all” approach is generally taken regarding how an individual will respond to affective computing manipulations and system designs (Proctor & van Zandt, 1994).

Furthermore, human factors and HCI research has traditionally been divorced from individual difference factors (McNeese, 2003; Szalma, 2009). Under this approach, users are all considered to be similar and individual differences are typically not measured. However, there has been a recent push in these domains calling for the inclusion of personality and individual difference factors as legitimate concerns in both research and application (see Szalma, 2009, for a listing of related sources). One reason to suspect that individual differences could influence how affective computing impacts someone is that previous research has shown that individuals exhibiting some combination of particular personality traits may be better suited for certain tasks than others (Matthews, Deary, &
human–human interactions have been successfully replicated in HCl. This is evident in the areas of human–computer trust (Muir, 1988; Muir & Moray, 1996; Reeves & Nass, 1996), attributions (Johnson et al., 2008), emotion (Boehner et al., 2007), and virtual community building (Lin, 2010).

Before individual differences can be effectively utilized in the design of affective computing platforms, more needs to be understood about which individual difference factors are likely to have an impact on a user’s experience in using affectively driven applications. Thus, the current study sought to identify the personality traits that have the largest role on impacting affect change caused by a computer program, as well as examining the role of gender. Computer games are an area in which affective interfaces are being designed (Gilleade & Allanson, 2003; Gilleade & Dix, 2004). Although the impact of reinforcement on players’ affect within such games is already being researched (Chumbley & Griffiths, 2006), the current research innovatively examines the way that individual difference factors, particularly personality, modulate affect change within a gaming paradigm. The HCI architecture used in this study, playing a computer-based trivia game, was a hedonic paradigm that was intended to be fun and to increase positively associated affect while decreasing negatively associated affect. The present study also manipulated the nature of the feedback given to the participants. It was hypothesized that positive affective feedback would result in greater changes in affect than simple, neutral feedback (Hypothesis 1).

Through the use of personality measures, it was possible to look at the relationships between different personality traits and affect change. This study involved an untried computing and research paradigm and was thus exploratory in nature; therefore we chose not to make individual predictions for each of the 13 affect scales that were assessed, but rather chose to make nondirectional hypotheses about the global impact of individual difference factors across all the scales in a multivariable analysis. It was hypothesized that participants’ personality trait scores would significantly impact affect change when comparing pre- and postscores on an affect inventory based on previous individual difference findings (Matthews et al., 2003; Szalma, 2008, 2009; Hypothesis 2). Specifically, we hypothesized that participants’ neuroticism scores would influence affect change based on the previously found correlations between neuroticism and affect (Charles et al., 2001; Costa et al., 1980; Mroczek & Kolarz, 1998; Hypothesis 3).

Two alternative, theory-based, hypotheses were proposed regarding gender. Hypothesis 4 proposed that women would be impacted by positive affective feedback more so than men and was based on the gender differences hypothesis and the previously established relationship between neuroticism and gender, as well as change in affect. On the other hand, Hypothesis 5 was based on the gender-similarities hypothesis, under which we would not observe gender differences in response to feedback.

Gender is another individual difference that could influence the way individuals use computers. However, the apparent difference between genders has been debated in psychology for more than a century. One popular theory is the gender differences hypothesis, which argues that males and females are considerably different psychologically. Some of these differences have emerged in personality literature (Feingold, 1994) and even HCI. Two findings from the personality literature that have a bearing on computer use are that women generally have higher levels of neuroticism and are more likely to externalize locus of control than men (Hall, 1984). In addition, Johnson, Veltri, and Hornik (2008) found that women were more likely than men to ascribe responsibility to a computer, both with a traditional user interface and one that utilized social cues. The alternative view is the gender similarity hypothesis, which states that males and females are similar in most psychological and behavioral variables. Although the gender differences hypothesis typically dominates current research, several meta-analyses have shown considerable support for the gender similarities hypothesis and point out the small effect sizes obtained in most studies that find gender differences (Feingold, 1994; Hyde, 2005).

Although individual difference traits and affect have not been explored systematically within the field of HCI or the domain of affective computing, there is existing social psychology research that is of relevance to beginning this line of research. According to Watson and Clark (1992) and Charles, Reynolds, and Gatz (2001), individuals who scored higher in neuroticism had higher ratings of negative affect and were less likely to show decreases in negative affect. They also found that higher levels of extraversion provided a buffer against decreases in positive affect, which is further supported by the findings of Costa, McCrae, and Arenberg (1980) and Mroczek and Kolarz (1998), who also found a positive correlation between positive affect and extraversion.

There is some disagreement in the literature regarding the relationship between neuroticism and negative affect. Whereas the originators of the Positive and Negative Affect Schedule–Expanded Form (PANAS-X), Watson and Clark (1992), have not found a relationship between positive affect and neuroticism, there is evidence that one may exist (Costa et al., 1980; Mroczek & Kolarz, 1998). Although these studies were primarily interested in human–human interaction, many findings in human–human interactions have been successfully
1. METHOD

1.1. Participants

Participants were 76 undergraduate students (46 female, 30 male) from New Mexico State University who enrolled in the study to fulfill a course requirement. The mean age was 19.8 years (SD = 3.3 years, range = 18–44 years).

1.2. Materials

The experiment was run on a desktop computer using a standard keyboard and mouse. The display was a 22-in. LCD monitor set at a resolution of 1920 × 1080 pixels. The computer was equipped with Microsoft Windows 7 and used E-Prime 2.0 to take the participant through the experimental session, which included various surveys and the trivia game. The trivia game was designed such that it would not be difficult for participants to do well while remaining somewhat challenging. Questions for the trivia game were generated by the experimenters or retrieved from http://www.triviaplaying.com. (See Appendix D for sample queries and answers.)

Personality traits were measured using the 44-item Big Five Inventory (BFI; John, Donahue, & Kentle, 1991). The Big Five structural model of personality traits is arguably the most popular model of personality currently and has been validated across numerous instruments and observers (McRae & Costa, 1987). The model consists of five personality scales—Agreeableness, Extraversion, Neuroticism, Openness, and Conscientiousness—and has been shown to converge with the Big Three and Alternative Five structural models of personality (Zuckerman, Kuhlman, Joireman, Teta, & Kraft, 1993). Two additional, standard individual difference measures that impact social interaction were also assessed. Narcissism was measured using the 16-item Narcissistic Personality Inventory (NPI-16; Ames, Rose, & Anderson, 2006) and self-esteem was measured using the Rosenberg Self-Esteem Scale (SES; Rosenberg, 1989), both widely used tests in individual difference research and have been shown to have a relationship with locus of control (Anderson, 1977; Biondo & MacDonald, 1971; Hjelle & Closer, 1970; Judge, Erez, Bono, & Thoresen, 2002), a factor that impacts HCI (Chak & Leung, 2004; Woodrow, 1990).

Participants’ affect before and after the trivia game was measured using the Positive and Negative Affect Schedule–Expanded Form (PANAS-X; Watson & Clark, 1994). The PANAS-X divides affect into 13 scales that belong to four categories: general dimension scales (negative and positive affect), basic negative emotion scales (fear, hostility, guilt, sadness), basic positive emotion scales (joviality, self-assurance, attentiveness), and other affective scales (shyness, fatigue, serenity, and surprise). Current-state affect before and after the trivia game was of interest, so questions in the PANAS-X took on the temporal format of “right now” (e.g., “Indicate to what extent you feel this way right now”).

1.3. Procedure

After giving consent, participants were seated at a computer in a quiet room and completed the NPI-16, Rosenberg SES, BFI, and PANAS-X (pre), in this order. Responses were made using the number keys on the keyboard. After filling out the measures, participants were shown the instructions for the trivia game on the screen.

The instructions stated that they would be playing a trivia game that consisted of general knowledge questions and they would accumulate virtual money for correct answers. Their goal was to accumulate as much virtual money as possible. Questions had variable monetary value and, to minimize the discomfort felt when one got a question wrong, participants were informed that they would not know how much a question was worth unless they answered it correctly. Questions were worth anywhere from $5 to $35, and the monetary value was arbitrarily assigned without regard to perceived difficulty.

Next, participants were informed about the nature of the computer’s feedback. The instructions they received depended on which condition they were in. In the neutral feedback condition, participants were told that after every question they would be informed whether their answer was right or wrong, how much money they earned from that answer (if it was correct), and how much money they had accumulated thus far. Verbal statements in the neutral feedback condition consisted solely of “that answer is correct” for correct answers and “that answer is incorrect” for incorrect answers.

In the positive feedback condition, participants were told that feedback from the computer would vary based on how well they were performing in the trivia game. They were told that doing well would yield positive feedback from the computer (meant to increase positively associated affect and decrease negatively associated affect), whereas poor performance would result in negative feedback. Participants were also told that performance would be evaluated based on the total accumulated money, not how many questions they got right. Unbeknownst to the participants, all feedback was positive, regardless of their performance in the game, again to increase positive emotions as a result of playing the game. As in the neutral condition, they also received feedback on how much money they had accumulated so far and, in the case of a correct answer, how much virtual money they had earned from that question. This condition was created to evaluate a common technique used in affective tutors that is intended to engage the student with their learning material to a greater degree by providing positive feedback (Burleson, 2006; Picard, 1997, 2003).

Feedback comments (see Appendix E) in the positive feedback condition were created by the experimenters and each assigned to specific questions. Some of the feedback statements were specific to a certain question and appeared only once (e.g., “Well done, that was the highest money value question in the whole game!” for a correct answer, or “That’s ok, this was the most missed question” for an incorrect answer). Some feedback was not specific and could appear multiple times.
(e.g., “Great job, you’re doing well on this task!” for a correct answer, or “Good try, you’re still doing well overall” for an incorrect answer). In the positive feedback condition, there were 49 unique feedback responses for correct answers and 46 unique feedback responses for incorrect answers. This helped maintain the illusion that the computer really was providing feedback based on the actual success of the participant.

Following reading the instructions, participants began the trivia game. Seventy multiple-choice questions were presented, one at a time; each question had four possible answers. Participants indicated their answer by clicking on a checkbox next to the answer. Participants were able to select only one answer choice per question and would click “Next” once they made their decision. To ensure that the positive feedback was believable, the first 10 questions were the same for all participants and were chosen because they were easy, in the hopes that participants got off to a good start. The next 50 questions were randomized. The final 10 questions were not randomized and designed so that the feedback in the positive feedback condition would tell the participant that they had accumulated a lot of money and were doing better than most participants had done in the game. These verbal prompts always occurred at the end of the game, again in the hopes of making the feedback more believable.

After playing the trivia game, participants filled out the PANAS-X (post) again. The entire study took approximately 30 min to complete. Participants were randomly assigned to their experimental condition (either neutral, \( n = 39 \), or positive, affective feedback, \( n = 37 \)) when they arrived for the study.

1.4. Design

The design of the study was experimental and involved two factors: the nature of the feedback, which was manipulated between subjects, and gender, which served as a subject variable. Participants in both conditions completed identical inventories. The PANAS-X instrument was administered before and after playing the trivia game. These pre- and postscores were used to compute affect change scores, which served as the dependent variables.

A number of analyses including multivariable and univariate analyses of variance (MANOVAs and ANOVAs), paired-samples \( t \) tests, and correlations were utilized. For the ANOVAs, the dependent variables were the absolute changes in affect scores for each of the 13 PANAS-X scales. Individual difference measures (age and scores on the NPI-16, Rosenberg SES, and BFI) served as covariates in the ANOVAs and as factors in the correlation analyses.

2. RESULTS

Participants’ change in affect was calculated separately for each of the 13 affect scales defined by the PANAS-X by subtracting participants’ pre-PANAS-X scores from their post-PANAS-X scores (see Table 1). As a check of the efficacy of

| TABLE 1 |
| Mean Affect Scores (Pre, Post, and Change) and Their Standard Deviations Across the 13 Scales of the Positive and Negative Affect Schedule–Expanded Form Separated by Feedback Condition |

<table>
<thead>
<tr>
<th>Affect Scale</th>
<th>Neutral Pre</th>
<th>Neutral Post</th>
<th>Neutral Change</th>
<th>Positive Change</th>
<th>Positive Pre</th>
<th>Positive Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative</td>
<td>14.5 (6.1)</td>
<td>12.9 (4.1)</td>
<td>-1.6</td>
<td>-2.1</td>
<td>14.2 (5.0)</td>
<td>12.1 (2.7)</td>
</tr>
<tr>
<td>Positive</td>
<td>30.2 (9.4)</td>
<td>30.7 (10.2)</td>
<td>0.5</td>
<td>2.0</td>
<td>29.8 (8.5)</td>
<td>31.8 (9.1)</td>
</tr>
<tr>
<td>Fear</td>
<td>8.5 (3.8)</td>
<td>7.3 (1.9)</td>
<td>-1.2</td>
<td>-1.3</td>
<td>8.7 (3.7)</td>
<td>7.4 (2.1)</td>
</tr>
<tr>
<td>Hostility</td>
<td>7.9 (2.7)</td>
<td>7.3 (2.3)</td>
<td>-0.6</td>
<td>-1.9</td>
<td>8.5 (3.6)</td>
<td>6.6 (1.2)</td>
</tr>
<tr>
<td>Guilt</td>
<td>8.1 (3.6)</td>
<td>6.9 (2.0)</td>
<td>-1.2</td>
<td>-0.9</td>
<td>7.4 (2.0)</td>
<td>6.5 (1.2)</td>
</tr>
<tr>
<td>Sadness</td>
<td>7.9 (3.7)</td>
<td>6.9 (2.7)</td>
<td>-1.0</td>
<td>-1.3</td>
<td>7.7 (2.8)</td>
<td>6.4 (2.2)</td>
</tr>
<tr>
<td>Joviality</td>
<td>23.1 (8.5)</td>
<td>23.9 (8.8)</td>
<td>0.8</td>
<td>2.5</td>
<td>23.5 (7.1)</td>
<td>26.0 (7.5)</td>
</tr>
<tr>
<td>Self-Assurance</td>
<td>17.4 (6.0)</td>
<td>16.8 (6.2)</td>
<td>-0.6</td>
<td>0.7</td>
<td>17.0 (5.3)</td>
<td>17.7 (5.7)</td>
</tr>
<tr>
<td>Attentiveness</td>
<td>13.4 (3.6)</td>
<td>13.2 (3.9)</td>
<td>-0.2</td>
<td>0.2</td>
<td>12.8 (3.5)</td>
<td>12.8 (3.7)</td>
</tr>
<tr>
<td>Shyness</td>
<td>6.4 (2.5)</td>
<td>5.2 (2.0)</td>
<td>-1.2</td>
<td>-1.2</td>
<td>6.3 (2.5)</td>
<td>5.1 (1.6)</td>
</tr>
<tr>
<td>Fatigue</td>
<td>12.0 (4.2)</td>
<td>9.7 (4.2)</td>
<td>-2.3</td>
<td>-2.7</td>
<td>10.9 (3.9)</td>
<td>8.2 (3.5)</td>
</tr>
<tr>
<td>Serenity</td>
<td>11.1 (2.9)</td>
<td>10.5 (3.0)</td>
<td>-0.6</td>
<td>-1.2</td>
<td>10.6 (2.4)</td>
<td>9.4 (2.6)</td>
</tr>
<tr>
<td>Surprise</td>
<td>5.5 (2.7)</td>
<td>6.6 (3.5)</td>
<td>1.1</td>
<td>2.2</td>
<td>5.6 (2.7)</td>
<td>7.8 (3.1)</td>
</tr>
</tbody>
</table>

*Note.* Standard deviations are in parentheses. Positive values indicate an increase and negative values indicate a decrease in a specific measure, regardless of the valence of the items. The change columns are displayed next to each other in order to facilitate comparison.
the hedonic computing paradigm, we examined how participants’ positive and negative affect was impacted as a result of playing the game. Participants showed a significant decrease in negative affect as a consequence of playing the game, \( \tau(75) = 4.19, p = .00008, \) Cohen’s \( d = 0.55 \) (corrected for correlation). Although participants did show increases in positive affect, these differences only approached marginal significance in a paired samples \( t \) test, \( \tau(75) = 1.57, p = .12, \) Cohen’s \( d = 0.18 \) (corrected for correlation). As illustrated in Table 1, the mean positive affect change in the positive feedback condition was actually four times greater than the change in the neutral feedback condition—not nearly the contrast observed with negative affect, yet the change in positive affect only approached marginal significance. This disparity is likely due the fact that the distribution of the positive affect scores had a mean and standard deviation that was more than twice than what was observed with negative affect. Taken together, the results provide evidence that the nature of the computing interaction was, in fact, hedonic, in the sense that participants’ positive affect generally rose and negative affect fell as a result of playing the game, regardless of the nature of the feedback.

Having completed the manipulation checks regarding the two general dimension scales (positive and negative affect) in the aforementioned analyses, the absolute values of the change along each scale were utilized in the following analyses because the PANAS-X contains affect scales that are of opposite valence and four of ambiguous valence. A MANOVA was performed on the entire dataset with absolute affect difference scores along the 13 PANAS-X scales serving as dependent variables. Feedback condition was the independent variable, gender was a fixed subject variable, and all of the individual difference factors were included as covariates. As summarized in Table 2, the analysis revealed a significant interaction between gender and feedback condition, Wilks’s \( \lambda = 0.65, F(13, 52) = 2.13, p = .02, \eta^2_p = 0.35, \) lending partial support to Hypothesis 1. Furthermore, in support of Hypotheses 2 and 3, neuroticism scores covaried with global affect change, Wilks’s \( \lambda = 0.68, F(13, 52) = 1.87, p = .056, \eta^2_p = 0.32, \) although the level of statistical significance was marginal. Three of the other individual difference factors also approached significant covariability with affect change: narcissism, Wilks’s \( \lambda = 0.7, F(13, 52) = 1.7, p = .09, \eta^2_p = 0.30; \) self-esteem, Wilks’s \( \lambda = 0.69, F(13, 52) = 1.78, p = .07, \eta^2_p = 0.31; \) and extraversion, Wilks’s \( \lambda = 0.71, F(13, 52) = 1.68, p = .09, \eta^2_p = 0.30. \) These four factors were retained in the consequent, gender-specific MANOVAs that explored the interaction between gender and feedback condition. Age and the remaining personality trait factors did not show a significant relationship with affect change and were not included in the following analyses.

The MANOVA of female participants’ affect change revealed a significant effect of feedback condition, Wilks’s \( \lambda = 0.51, F(13, 28) = 2.09, p = .05, \eta^2_p = 0.49, \) and a significant contribution of participants’ neuroticism scores, Wilks’s \( \lambda = 0.44, F(13, 28) = 2.73, p = .01, \eta^2_p = 0.56. \) Univariate analyses revealed that feedback condition had a significant impact on positive affect change, \( F(1, 40) = 5.66, p = .02, \eta^2_p = 0.12, \) with women in the positive feedback reporting greater changes in positive affect scores (\( M_A = 3.57, \) and a marginally significant impact on the changes in two of the three basic positive affect scales: joviality (\( M_A = 1.98, \) \( F(1, 40) = 3.32, p = .08, \eta^2_p = 0.08, \) and attentiveness (\( M_A = 1.78, \) \( F(1, 40) = 4.035, p = .051, \eta^2_p = 0.09. \) As the means in Table 1 illustrate, the findings are consistent with Hypothesis 1. In addition, there was an unexpected, marginally significant difference in the degree to which participants were surprised (\( M_A = 1.23, \) \( F(1, 40) = 3.0, p = .09, \eta^2_p = 0.07. \) Condition did not significantly impact change in females’ affect along the other nine scales (\( ps > 0.1. \) Univariate analyses revealed that neuroticism scores significantly covaried with changes in negative affect, \( F(1, 40) = 8.8, p = .005, \eta^2_p = 0.18, \) and two of the four basic negative affect scales: fear, \( F(1, 40) = 6.13, p = .02, \eta^2_p = 0.13, \) and hostility, \( F(1, 40) = 5.76, p = .02, \eta^2_p = 0.13. \) There was also a marginal covariance of neuroticism scores with attentiveness scores, \( F(1, 40) = 3.88, p = .06, \eta^2_p = 0.09. \) To determine the magnitude and direction of the significant covariances in the aforementioned MANOVA, gender-specific correlations between the previous factors and neuroticism were conducted (the significance level was set to \( \alpha = 0.01 \) to control for Type I error). Neuroticism scores correlated significantly with change in negative affect, \( r(45) = -0.38, p = .008, \) and fear, \( r(45) = -0.36, p = .001. \) The strength of the correlations between neuroticism and fear, \( r(45) = -0.25, p = .09; \) hostility, \( r(45) = -0.26, p = .08; \) and attentiveness, \( r(45) = -0.02, p = .001, \) did not reach statistical significance.

The interaction between gender and condition was further explored by conducting a MANOVA of male participants’ affect change. This analysis revealed solely a significant contribution of narcissism scores (Wilks’s \( \lambda = 0.182, \) \( F(13, 12) = 4.15, p = .01, \eta^2_p = 0.82, \) on affect change. The impact of feedback condition, Wilks’s \( \lambda = 0.47, F(13, 12) = 1.04, p > .1, \) or any of the other covariates was not significant. Unlike the women’s results, the men were unaffected by the type of feedback they received, contrary to Hypothesis 1. Univariate analyses revealed that narcissism primarily impacted change in Serenity, \( F(1, 24) = 7.94, p = .01, \eta^2_p = 0.25. \) A post hoc correlation analysis revealed a marginal negative relationship between narcissism and change in serenity, \( r(29) = -0.35, p = .056. \) No other significant relationships were observed in the univariate ANOVAs among the individual difference factors and affect change.

To further examine whether the observed relationships between personality factors and affect were consistent with previously reported findings (e.g., Charles et al., 2001; Watson & Clark, 1992) exploratory post hoc correlation analyses were conducted among Big Five personality scales and the pre-, post-, and difference positive and negative affect scores. The significant relationships are summarized in Table 3. The significance level was set to \( \alpha = .005 \) to control for Type I error. Neuroticism correlated significantly with initial levels of
### Table 2

Results of Main MANOVA and Two Secondary MANOVAs That Were Utilized to Explore the Discovered Interaction Between Gender and Type of Feedback

<table>
<thead>
<tr>
<th>Factor</th>
<th>Wilks's $\lambda$</th>
<th>$F$</th>
<th>$df_{H}, df_{E}$</th>
<th>$p$</th>
<th>$\eta_{p}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main MANOVA of both genders (all factors included)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interaction between gender and feedback</td>
<td>0.65</td>
<td>2.13</td>
<td>13, 52</td>
<td>.03*</td>
<td>0.35</td>
</tr>
<tr>
<td>Feedback condition</td>
<td>0.77</td>
<td>1.20</td>
<td>13, 52</td>
<td>.30</td>
<td>0.23</td>
</tr>
<tr>
<td>Gender</td>
<td>0.79</td>
<td>1.06</td>
<td>13, 52</td>
<td>.42</td>
<td>0.21</td>
</tr>
<tr>
<td>Age</td>
<td>0.88</td>
<td>0.53</td>
<td>13, 52</td>
<td>.89</td>
<td>0.12</td>
</tr>
<tr>
<td>Openness</td>
<td>0.90</td>
<td>0.43</td>
<td>13, 52</td>
<td>.95</td>
<td>0.10</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>0.86</td>
<td>0.64</td>
<td>13, 52</td>
<td>.81</td>
<td>0.14</td>
</tr>
<tr>
<td>Extraversion</td>
<td>0.71</td>
<td>1.68</td>
<td>13, 52</td>
<td>.09¹</td>
<td>0.30</td>
</tr>
<tr>
<td>Agreeability</td>
<td>0.79</td>
<td>1.08</td>
<td>13, 52</td>
<td>.40</td>
<td>0.21</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>0.68</td>
<td>1.87</td>
<td>13, 52</td>
<td>.06¹</td>
<td>0.32</td>
</tr>
<tr>
<td>NPI</td>
<td>0.70</td>
<td>1.70</td>
<td>13, 52</td>
<td>.09¹</td>
<td>0.30</td>
</tr>
<tr>
<td>RES</td>
<td>0.69</td>
<td>1.78</td>
<td>13, 52</td>
<td>.07¹</td>
<td>0.31</td>
</tr>
<tr>
<td>MANOVA of women’s scores (reduced set of factors)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feedback condition</td>
<td>0.51</td>
<td>2.09</td>
<td>13, 28</td>
<td>.05*</td>
<td>0.49</td>
</tr>
<tr>
<td>Extraversion</td>
<td>0.68</td>
<td>1.02</td>
<td>13, 28</td>
<td>.46</td>
<td>0.32</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>0.44</td>
<td>2.73</td>
<td>13, 28</td>
<td>.01*</td>
<td>0.56</td>
</tr>
<tr>
<td>NPI</td>
<td>0.60</td>
<td>1.45</td>
<td>13, 28</td>
<td>.20</td>
<td>0.40</td>
</tr>
<tr>
<td>RES</td>
<td>0.62</td>
<td>1.31</td>
<td>13, 28</td>
<td>.27</td>
<td>0.38</td>
</tr>
<tr>
<td>MANOVA of men’s scores (reduced set of factors)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feedback condition</td>
<td>0.47</td>
<td>1.04</td>
<td>13, 12</td>
<td>.47</td>
<td>0.53</td>
</tr>
<tr>
<td>Extraversion</td>
<td>0.40</td>
<td>1.40</td>
<td>13, 12</td>
<td>.28</td>
<td>0.60</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>0.43</td>
<td>1.25</td>
<td>13, 12</td>
<td>.35</td>
<td>0.58</td>
</tr>
<tr>
<td>NPI</td>
<td>0.18</td>
<td>4.15</td>
<td>13, 12</td>
<td>.01*</td>
<td>0.82</td>
</tr>
<tr>
<td>RES</td>
<td>0.40</td>
<td>1.36</td>
<td>13, 12</td>
<td>.30</td>
<td>0.60</td>
</tr>
</tbody>
</table>

*Note. MANOVA = multivariate analysis of variance.

¹Marginal significance. *Significance at $\alpha = 0.05$.

None of the other personality traits were correlated with positive or negative affect change, but several were correlated with pre- or postlevels of affect. Self-esteem was significantly correlated with postgame levels, $r(74) = 0.43, p = .00009$, of positive affect. Similarly, narcissism was significantly correlated with posttrivia game levels of positive affect, $r(74) = 0.33, p = .004$, and marginally correlated with pregame positive affect, $r(74) = 0.31, p = .007$. Extraversion was significantly correlated, $r(74) = 0.367, p = .001$, with initial levels of positive affect. All other relationships failed to reach significance.

### 3. Discussion

As expected from a hedonic paradigm, participants’ affect was positively impacted by the structured, relatively easy trivia game. There was a marginal increase in positive affect and
a significant decrease in negative affect providing a check of the trivia game manipulation. The use of positive feedback did produce a significantly greater magnitude of affect change than did neutral feedback for women, but not for men; thus our first hypothesis was only partially supported. Hypotheses 2, 3, and 4 were supported; the results showed that individual differences, including gender and personality trait factors like neuroticism, extraversion, narcissism, and self-esteem can impact the degree to which affect changes as an individual uses a computer. These results are consistent with previously established personal computing differences stemming from factors such as age and gender, which have been shown to impact the dynamic of the HCI (Chou & Chen, 2009; Lavie & Meyer, 2010). In addition, when the positive affective feedback provided in this paradigm is viewed as a behavioral reinforcer, these data are consistent with the results of Chumbley and Griffiths (2009), who showed that reinforcement contingencies can significantly impact players’ affect.

The differential response in men and women to affective feedback supports the gender differences hypothesis. Feingold (1994) and Hyde (2005) asserted that the effect sizes for gender differences reported by Hall (1984) and other studies supporting the gender differences hypothesis were generally small, yet in the current study the effect size of the interaction between gender and feedback condition was large ($\eta_p^2 = 0.35$), using Cohen’s rule of thumb (Cohen, 1988). In the exploration of the interaction between gender and feedback condition, it was discovered that neuroticism measures covaried significantly with female participants’ affect change scores but not with male participants’ scores. Based on the previously established relationship between neuroticism and affect change (Charles et al., 2001; Watson & Clark, 1992), the gender difference in neuroticism found in the current study could have contributed to the differential response to affective feedback between men and women.

We replicated the findings of Watson and Clark (1992) and Charles et al. (2001) by finding a positive correlation between neuroticism and pre- and postlevels of negative affect and a negative correlation between neuroticism and the change in negative affect. Additional support for the impact of personality on affect was the observed correlations between initial levels of positive affect and extraversion (positive), which was also consistent with the aforementioned studies. There was also a correlation between positive affect and neuroticism (negative), which was consistent with the findings of Costa et al. (1980) and Mroczek and Kolarz (1998) but contrary to what was found by Watson and Clark (1992).

These results provide evidence that individual differences should be an important part of further HCI and human factors research, as well as outline which personality traits have the potential to significantly impact users’ affect as a result of interfacing with a computer. A user’s gender and neuroticism or narcissism level, depending on the gender, should be a primary focus in computer programs that are designed specifically to manipulate affect. Our results suggest that affective computing manipulations like providing positive feedback may be generally ineffective for men but may work for female users, particularly those with high neuroticism levels.

Knowing the effects of such factors thus opens the possibility of fitting custom interfaces to users, for example, providing training for the individuals in areas where they are less naturally suited or using computerized teaching tailored to specific individuals. Similarly, in considering individual differences in affective computing, computers would ideally be able to adapt to each user according to the individual’s personality traits and learning styles. This would allow computers to take advantage of the user’s strengths and adjust to compensate for the users’ potential shortcomings. Such adaptive features would be useful in both education and training, for example, adjusting a tutoring session to increase interest if the computer detects student boredom (Carberry & de Rosis, 2008) or tailoring the degree to which feedback is affective based on current and future findings. Furthermore, these results show that tutoring platforms utilizing affective methods in education, teaching, and training (D’Mello, Craig, Witherspoon, McDaniel, & Graesser, 2008; McQuiggan, Mott, & Lester, 2008) and in everyday settings (Neviarouskaya

<table>
<thead>
<tr>
<th>Individual Difference Measure</th>
<th>Affect Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-Positive</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>$-0.368$</td>
</tr>
<tr>
<td>Self-Esteem</td>
<td>$0.400$</td>
</tr>
<tr>
<td>Narcissism</td>
<td>$0.310^*$</td>
</tr>
<tr>
<td>Extraversion</td>
<td>$0.310^*$</td>
</tr>
</tbody>
</table>

Note. Significance was defined as $\alpha = 0.005$ to control for Type I error.

$^*$Marginal significance; all other correlations in the table are statistically significant.
et al., 2010) could have vastly different effects depending on the user’s personality traits.

Apart from the current practical applications such as tutoring, training, and education, individual differences may be applied hedonically, or the branch of human factors that promotes pleasurable human–technology interaction (Hancock, Pepe, & Murphy, 2005). The extension of HCI aims to create an enjoyable experience for the user, going beyond just satisfying users’ needs of safety, functionality, and usability. This is done by promoting user satisfaction in the design of the system and allowing for individuation so that the user can achieve “personal perfection” with the system by customizing it to fit their personal preferences and traits. Based on the current results, it is possible to envision a system that accounts for individual differences that is capable of specifically targeting each person’s own personal interaction style, leading to increased pleasure for the user.

The current study was a simple, hedonic computing scenario intended to increase positive affect while decreasing negative affect by having people play a simple trivia game, while also providing positive feedback to half of the participants and neutral feedback to the others. Participants’ affect did improve as a consequence of playing the game, and the positive feedback modulated the degree to which affect changed for women, but not for men. In addition, it was observed that personality trait factors like neuroticism, narcissism, extraversion, and self-esteem impacted the degree of the affect change, again differentially for the genders. More generally, the results indicate the significant impact that individual differences can make in affective HCI. They also motivate a two-headed user-centered design and affective computing research effort, the traditional approach that involves sensing users’ affective states, as well as the approach proposed here, in which individual differences are used to create adaptive interfaces that are can be tailored individually to users.

REFERENCES


Boehner, K., DePaula, R., Dourish, P., & Sengers, P. (2007). How emotion is used to create adaptive interfaces that are can be tailored individually to users.


---

**ABOUT THE AUTHORS**

Jeremy D. Schwark earned his B.A. in Psychology at Trinity University and his M.A. in Experimental Psychology at New Mexico State University, where he is currently a Ph.D. student in Applied Cognitive Psychology. He is a HHMI Scientific Teaching Fellow and research associate in Dr. Igor Dolgov’s PACMANe Lab.

Igor Dolgov earned a BSE in Computer Science from Princeton University and a Ph.D. in Psychology-Arts, Media, & Engineering from ASU via NSF’s IGERT Program. An assistant professor at NMSU, he is the lead investigator at The Perception, Action, and Cognition, in Mediated, Artificial, and Natural Environments (PACMANe/*"pacman"*) laboratory.

Daniel Hor has a B.S. in Computer Science from Michigan Technological University and is currently completing his M.A. in Experimental Psychology at New Mexico State University. He specializes in HCI, usability, and other Human Factors research at the PACMANe lab, and has prior work experience in publications design and editing.

William Graves earned a B.A. in Psychology from Wichita State University and a M.A. in Experimental Psychology from New Mexico State University. He is a program specialist at 21st Century Aerospace with a focus on unmanned aircraft systems. His research interests include the cognitive and social effects of video games.

**APPENDIX A**

**Rosenberg Self-Esteem Scale**

Answers: Strongly Agree, Agree, Disagree, or Strongly Disagree

1. I feel that I am a person of worth, at least on an equal plane with others.
2. I feel that I have a number of good qualities.
3. All in all, I am inclined to feel that I am a failure.
4. I am able to do things as well as most other people.
5. I feel I do not have much to be proud of.
6. I take a positive attitude towards myself.
7. On the whole, I am satisfied with myself.
8. I wish I could have more respect for myself.
9. I certainly feel useless at times.
10. At times I think I am no good at all.

**APPENDIX B**

**Narcissistic Personality Inventory – 16**

For each pairing, choose the one that closest matches your feelings or beliefs.

1. ____ I really like to be the center of attention
   ____ It makes me uncomfortable to be the center of attention
3. Does a thorough job.
4. Is depressed, blue.
5. Is original, comes up with new ideas.
6. Is reserved.
7. Is helpful and unselfish with others.
8. Can be somewhat careless.
10. Is curious about many different things.
11. Is full of energy.
12. Starts quarrels with others.
13. Is a reliable worker.
14. Can be tense.
15. Is ingenious, a deep thinker.
16. Generates a lot of enthusiasm.
17. Has a forgiving nature.
18. Tends to be disorganized.
19. Worries a lot.
20. Has an active imagination.
21. Tends to be quiet.
22. Is generally trusting.
23. Tends to be lazy.
24. Is emotionally stable, not easily upset.
25. Is inventive.
26. Has an assertive personality.
27. Can be cold and aloof.
28. Perseveres until the task is finished.
29. Can be moody.
30. Values artistic, aesthetic experiences.
31. Is sometimes shy, inhibited.
32. Is considerate and kind to almost everyone.
33. Does things efficiently.
34. Remains calm in tense situations.
35. Prefers work that is routine.
36. Is outgoing, sociable.
37. Is sometimes rude to others.
38. Makes plans and follows through with them.
40. Likes to reflect, play with ideas.
41. Has few artistic interests.
42. Likes to cooperate with others.
43. Is easily distracted.
44. Is sophisticated in art, music, and literature.

APPENDIX D

Sample multiple-choice questions from the trivia game. Each had four possible responses. Each sentence indicates a single trial (even if there are multiple on a line).

What does the F in FBI stand for? Which cartoon character says “What’s up, Doc?”
What sport did Michael Jordan play? What color are the stars in the United States’ flag?
In football, what city are the Cowboys from? What color is the M in McDonald’s?
The sale of what was prohibited in America during prohibition?
Where is the city of Dublin?
Which John was the star of the films *Grease* and *Pulp Fiction*?
San Diego is in which state?
San Antonio International Airport is located in which state?
Which city was Marco Polo Airport built? Which state is called the Bluegrass State?
Who was president during World War I? Which company made *The Lion King* CD ROM?
What type of pills is known in the pharmacy biz as “pillows”?
How many legs do insects have?
Alcatraz was a prison on an island off the coast of which US city? What’s the hardest mineral?
In mythology, Pegasus was a horse with what? In what movie did R2-D2 and C-3PO star in?
What is a mammal with a pouch for its young called? What is the best-selling book of all time?
In which country is an Afghani a unit of currency? What is the main color of the UN flag?
How was the Cote d’Ivoire previously known? Buckingham Palace is in which English City?
In *Peanuts*, what color is Woodstock the bird? What planet is closest in size to our moon?
What is the name of Sherlock Holmes’s sidekick? Where is the Yankees baseball team from?
What country was the first to send a man into space? Portugal lies east of which ocean?
What is the longest river in the world? What was developed in the 40s Manhattan Project?

**APPENDIX E**

Examples of the kind of feedback participants received for getting questions right and wrong. Each sentence indicates a single instance of feedback (even if there are multiple on a line).

Positive feedback when getting a question correct:

You’re off to a good start! Well done! Keep it up! Well done, you sure know your history!
Great job, that was a tough question! Good job! So far, you’re doing better than most people!
Amazing, almost no one gets that question right! Great job, that was worth a lot of money!
Well done, you’re doing well on the geography questions! Excellent, keep it up!
Very good, you’re doing better on this task than most participants! Excellent work!
Wow, this was the most missed question and you got it right! Correct answer!
Well done, that question was worth a lot of money! Nice going, you know your mythology!

Positive feedback when getting a question incorrect:

Keep trying. Almost got it. Your answer choice was close. Nice try.
You’re still doing OK. Good try. Good try, that was a hard one. Almost, you were on the right track. Nice try, eating fast food is bad for you anyway.
Good try, that was a hard question. You were on the right track.
Almost, you had the right era. Good try, you’re still doing well. Almost, not many people got this right. That’s OK, this was the most missed question.
Nice try, it wasn’t worth much money anyway. Nice try, this was a hard one.
Good try, you’re still doing very well. Good try, that question was deceptively hard.
Close. Nice try, most people missed that one as well. Close, that book is high selling as well.