HUMAN–COMPUTER INTERACTION, 2007, Volume 22, pp. 325–353

Virtual Interpersonal Touch: Expressing and Recognizing Emotions Through Haptic Devices

Jeremy N. Bailenson, Nick Yee, Scott Brave, Dan Merget, and **David Koslow**
Stanford University

ABSTRACT

This article examines the phenomenon of Virtual Interpersonal Touch (VIT), people touching one another via force-feedback haptic devices. As collaborative virtual environments become utilized more effectively, it is only natural that interactants will have the ability to touch one another. In the work presented here, we used relatively basic devices to begin to explore the expression of emotion through VIT. In Experiment 1, participants utilized a 2 DOF force-feedback joystick to express seven emotions. We examined various dimensions of the forces generated and subjective ratings of the difficulty of expressing those emotions. In Experiment 2, a separate group of participants attempted to recognize the record-

Jeremy Bailenson is a social scientist with an interest in digital human representation; he is an Assistant Professor in the Department of Communication at Stanford University and is the founding director of the Virtual Human Interaction Lab. **Nick Yee** is a social scientist with an interest in self-representation and social interaction in virtual worlds; he is a Ph.D. candidate in the Department of Communication at Stanford University. **Scott Brave** is an HCI practitioner with an interest in the social aspects of computing; he received his Ph.D. from Stanford University and is now the founder and chief scientist at Baynote, Inc. **Dan Merget** is a computer scientist with an interest in graphics; he received his Masters in Computer Science from Stanford University. **David Koslow** is a computer scientist with an interest in virtual human interaction; he is an undergraduate student in Computer Science at Stanford University.

CONTENTS

ings of emotions generated in Experiment 1. In Experiment 3, pairs of participants attempted to communicate the seven emotions using physical handshakes. Results indicated that humans were above chance when recognizing emotions via

VIT but not as accurate as people expressing emotions through nonmediated handshakes. We discuss a theoretical framework for understanding emotions expressed through touch as well as the implications of the current findings for the utilization of VIT in human–computer interaction.

1. INTRODUCTION

There are many reasons to support the development of collaborative virtual environments (Lanier, 2001). One major criticism of collaborative virtual environments, however, is that they do not provide emotional warmth and nonverbal intimacy (Mehrabian, 1967; Sproull & Kiesler, 1986). In the work presented here, we empirically explore the augmentation of collaborative virtual environments with simple networked haptic devices to allow for the transmission of emotion through *virtual interpersonal touch* (VIT).

1.1. Emotion in Social Interaction

Interpersonal communication is largely nonverbal (Argyle, 1988), and one of the primary purposes of nonverbal behavior is to communicate subtleties of emotional states between individuals. Clearly, if social interaction mediated by virtual reality and other digital communication systems is to be successful, it will be necessary to allow for a full range of emotional expressions via a number of communication channels. In face-to-face communication, we express emotion primarily through facial expressions, voice, and touch. Although emotion is also communicated through other nonverbal gestures such as posture and hand signals (Cassell & Thorisson, 1999; Collier, 1985), in this article we focus on emotions transmitted via face, voice, and touch.

In a review of the emotion literature, Ortony and Turner (1990) discussed the concept of basic emotions. These fundamental emotions (e.g., fear) are the building blocks of other more complex emotions (e.g., jealousy). Furthermore, many people argue that these emotions are innate and universal across cultures (Plutchik, 2001). In terms of defining the set of basic emotions, previous work has provided very disparate sets of such emotions. For example, Watson (1930) limited his list to "hardwired" emotions such as fear, love, and rage. On the other hand, Ekman and Friesen (1975) limited their list to those discernable through facial movements such as anger, disgust, fear, joy, sadness, and surprise.

The psychophysiology literature adds to our understanding of emotions by suggesting a fundamental biphasic model (Bradley, 2000). In other words, emotions can be thought of as variations on two axes—hedonic valence and intensity. Pleasurable emotions have high hedonic valences, whereas negative emotions have low hedonic valences. This line of research suggests that although emotions may appear complex, much of the variation may nonetheless be mapped onto a two-dimensional scale. This notion also dovetails with research in embodied cognition that has shown that human language is spatially organized (Richardson, Spivey, Edelman, & Naples, 2001). For example, certain words are judged to be more "horizontal," whereas other words are judged to be more "vertical."

In the work presented here, we were not concerned predominantly with what constitutes a basic or universal emotion. Instead, we attempted to identify emotions that could be transmitted through virtual touch and provide an initial framework for classifying and interpreting those digital haptic emotions. To this end, we reviewed theoretical frameworks that have attempted to accomplish this goal with other nonverbal behaviors—most notably, facial expressions and paralinguistics.

1.2. Facial Expressions

Research in facial expressions has received much attention from social scientists for the past 50 years. Some researchers argue that the face is a portal to one's internal mental state (Ekman & Friesen, 1978; Izard, 1971). These scholars argue that when an emotion occurs, a series of biological events follow that produce changes in a person; one of those manifestations is movement in facial muscles. Moreover, these changes in facial expressions are also correlated with other physiological changes such as heart rate or blood pressure (Ekman & Friesen, 1976). Alternatively, other researchers argue that the correspondence of facial expressions to actual emotion is not as high as many think. For example, Fridlund (1994) believes that people use facial expressions as a tool to strategically elicit behaviors from others or to accomplish social goals in interaction. Similarly, other researchers argue that not all emotions have corresponding facial expressions (Cacioppo, Bernston, Klein, & Poehlmann, 1997). Nonetheless, most scholars would agree that there is some value to examining facial expressions of another if one's goal is to gain an understanding of that person's current mental state.

Ekman's groundbreaking work on emotions has provided tools to begin forming dimensions on which to classify his set of six basic emotions (Ekman & Friesen, 1975). Figure 1 provides a framework for the facial classifications developed by those scholars.

There has recently been a great surge of work to develop automatic algorithms to identify emotional states from a video image of facial movements. Early work developed a facial action coding system in which coders manually identified anchor points on the face in static images (Ekman & Friesen, 1978). Similarly, computer scientists have developed vision algorithms that automatically find similar anchor points with varying amounts of success (see Essa & Pentland, 1994, for an early example). As computer vision algorithms and perceptual interfaces become more elegant (see Turk & Kölsch, 2004, for a re-

Figure 1. **Characteristics of six emotions discernable through facial expressions.**

Surprise: brows raised, eyelids opened and more of the white of the eye is visible, jaw drops open without tension or stretching of the mouth
Fear: brows raised and drawn together, forehead wrinkles drawn to the center, mouth is open, lips are slightly tense or stretched and drawn back
Disgust: upper lip is raised, lower lip is raised and pushed up to upper lip or it is lowered, nose is wrinkled, cheeks are raised, lines below the lower lid, brows are lowered
Anger: brows lowered and drawn together; vertical lines appear between brows; lower lid is tensed and may or may not be raised; upper lid is tense and may or may not be lowered due to brows' action; eyes have a hard stare and may have a bulging appearance; lips are either pressed firmly together with corners straight or down or open, tensed in a squarish shape; nostrils may be dilated (could occur in sadness too) unambiguous only if registered in all three facial areas
Joy: corners of lips are drawn back and up, mouth may or may not be parted with teeth exposed or not, a wrinkle runs down from the nose to the outer edge beyond lip corners, cheeks are raised, lower eyelid shows wrinkles below it and may be raised but not tense, crow's-feet wrinkles go outward from the outer corners of the eyes.
Sadness: inner corners of eyebrows are drawn up, skin below the eyebrow is triangulated with inner corner up, upper lid inner corner is raised, corners of the lips are drawn or lip is trembling

view), it is becoming possible to measure the emotional state of people in real time, based on algorithms that automatically detect facial anchor points, and then categorize those points into emotions that have been previously identified using some type of learning algorithm. These systems sometimes attempt to recognize specific emotions (Michel & El Kaliouby, 2003) or alternatively attempt to gauge binary states such as general affect (Picard & Bryant Daily, 2005). In our work we attempt to accomplish a similar goal with expression of emotions through touch.

1.3. Voice

Nass and Brave (2005) provided a thorough review of the literature on voice and emotion. In terms of inferring aspects of emotions from vocal communication, arousal is the most readily discernible feature, but voice can also provide indications of valence and specific emotions through acoustic properties such as pitch range, rhythm, and amplitude or duration changes (Ball & Breese, 2000; Scherer, 1989). A bored or sad user, for example, will typically exhibit slower, lower pitched speech, with little high-frequency energy, whereas a user experiencing fear, anger, or joy will speak faster and louder, with strong high-frequency energy and more explicit enunciation (Picard, 1997). Murray and Arnott (1993) provided a detailed account of the vocal effects associated with several basic emotions.

1.4. Virtual Interpersonal Touch

In virtual reality, voice expression of emotion is easy through digitized audio streams. Facial expression is more challenging but certainly possible given recent advances in the computer vision tracking algorithms previously discussed. However, person-to-person haptic interaction, both because of the difficulty of constructing large force-feedback devices as well as the dearth of research in psychology on touching behavior (compared to other nonverbal behavior—see Argyle, 1988, for a review), has received less attention than face and voice.

We know that in general, touch tends to increase trust. For example, waiters who briefly touch their customers receive higher tips than those who do not (Crusco & Wetzel, 1984). In face-to face communication, people use touch to add sincerity/establish trust (valence), add weight/urgency, mark significance (arousal), and adhere to formalized greetings and parting gestures such as handshakes. However, touch is not used as often as facial expressions and voice intonation changes. Some reasons for this discrepancy are that touch is one-to-one only, not one-to-many as the other cues are. In other words, touch is inefficient. Furthermore, touch can be inconvenient and requires close distance and physical coupling (restriction of movement). Finally, touch may be overly intimate or socially inappropriate for many interactions (Burgoon & Walther, 1990), as touch is one of the most definitive markers of intimacy in social interaction.

Whereas handshaking is the most common social interaction that involves touch, very little empirical research has been done with regards to how handshaking relates to other variables, such as emotion. A notable exception is a study that investigated how variations in handshaking relate to personality and gender (Chaplin, Phillips, Brown, Clanton, & Stein, 2000). In that study, research assistants were trained to initiate a handshake with participants and rate the handshakes on a set of measures—completeness of grip, temperature, dryness, strength, duration, vigor, texture, and eye contact. Participants then filled out personality inventories. Substantial correlations among the handshaking measures led the researchers to create a composite, which they termed "firm handshake." Male participants were found to have firmer handshakes than female participants, and firmer handshakes were positively correlated with Extraversion and Openness to Experience on the Big-5 personality measures. One of the key contributions of the study was in demonstrating the link between personality and behavior and how personality might in fact be inferred from behavior. The goal of our studies is to demonstrate the ability to infer specific emotions from haptic behavior.

Previous work on virtual haptic communication and force-feedback has been largely used to simulate physical interaction between a human being and an inanimate object. However, there have been some projects designed to explore virtual interpersonal touch. One of the first attempts at multiuser force-feedback interaction, Telephonic Arm Wrestling (White & Back, 1986), provided a basic mechanism to simulate the feeling of arm wresting over a telephone line. Later on, Fogg, Cutler, Arnold, and Eisback (1998) described HandJive, a pair of linked handheld objects for playing haptic games. Similarly, InTouch (Brave, Ishii, & Dahley, 1998) is a desktop device that employs force-feedback to create the illusion of a shared physical object over distance, enabling simultaneous physical manipulation and interaction. Recently, Kim and colleagues (2004) developed haptic interaction platforms that allow multiple users to experience VIT without network delay. There have been other notable examples of projects geared toward allowing VIT (Chang, O'Modhrain, Jacob, Gunther, & Ishii, 2002; Clynes, 1977; Goldberg & Wallace, 1993; Noma & Miyasato, 1997; Oakley, Brewster, & Gray, 2000; Strong & Gaver, 1996). Many of these projects report positive reactions from users based on informal user testing.

Although there has been some work on the design side of VIT, very little is known about the psychological effects of haptic communication, although some research has begun to explore this issue. Basdogan, Ho, Slater, and Shrinivasan (1998) ran a series of studies in which participants used haptic devices to perform a collaborative task and could feel the digital avatars of one another while performing the task. Their results demonstrated that adding VIT to a visual interaction improved performance on a spatial task and increased subjective ratings of "togetherness" (see also Sallnas, Rassmus-Grohn, & Sjostrom, 2000). A study by Brave, Nass, and Sirinian (2001) presented participants with a screen-based maze. Participants were trying to either compete or cooperate with an alleged other player, and they either received haptic feedback or visual feedback from the other alleged player. Their results demonstrated that VIT caused changes in trust among the players; in competitive tasks, VIT increased subjective ratings of trust, whereas in cooperative tasks VIT decreased ratings of trust.

The results from these two studies examining VIT in user studies are extremely encouraging. VIT substantially changes an interaction, both in terms of task performance and subjective emotions toward other participants. Haptic communication has potential, because we know that the phenomenon of touching another human being is powerful but largely unused in virtual environments. VIT is uniquely compelling because we can use VIT to accomplish transformed social interaction (Bailenson, Beall, Loomis, Blascovich, & Turk, 2004). Transformed social interaction allows people in immersive virtual environments to accomplish nonverbal behaviors, appearances, and other interaction skills that are not possible in the physical world by allowing a strategic decoupling between rendered and performed behaviors.

In other words, with VIT we can create transformed haptic communication scenarios that are not possible in the physical world. For example, we can scale up or down aspects of the force behind VIT to accomplish interaction goals that are more appropriate to a given social context. Moreover, we can accomplish one-to-many interactions, allowing for a haptic gesture to be received by dozens of people at once. Finally, communication of emotion in virtual reality does not necessarily have to copy the real world; instead it can be abstracted (Brewster & Brown, 2004). We have the opportunity to explore alternate channels of emotional communication (e.g., avatars that change color when touched or using the facial expressions of Person A to regulate the degree to which Person B receives haptic feedback in a handshake).

As pointed out by a number of researchers (Durlach & Slater, 2000; Hansson & Skogg, 2001; Pantic & Rothkrantz, 2003; Rovers & Essen, 2004), it is essential to begin to develop a framework for understanding emotions communicated through haptic devices. For example, it has been shown that gesture recognition over video streams enhances remote collaboration (Fussell et al., 2004). Building haptic devices that recognize and could generate emotions would further enhance this remote collaboration paradigm. The goal of our work is to provide a set of studies that begins to test such a framework.

1.5. Deliberately Expressed Emotions Versus Automatic Leakage

One theme previously discussed in regards to facial expressions is the distinction between actively creating an emotional behavior for a strategic goal (deliberate), compared to an uncontrollable response to an emotion that is expressed without the person being able to control the behavior that controls the emotion (automatic). For example, there is research by Paul Ekman and colleagues on "The Duchenne Smile" (Ekman, Davidson, & Friesen, 1990), which is a specific and automatic type of smile that correlates with other physiological and behavioral predictors of actual enjoyment. Smiles that are more deliberate are qualitatively different than the automatic smiles and tend to have mouth movements that are similar to genuine smiles but fewer eye movements. In sum, some emotional facial expressions are deliberate, whereas others are automatic, and in terms of facial expressions and voice, it is possible to reliably differentiate the two. Indeed, there is a huge amount of research attempting to detect deception through facial and vocal cues (see Ekman, 2001, for a review).

As Nass and Brave (2005) pointed out, much research studying emotion in human–computer interaction (HCI) is problematic because it tends to examine photographs, voices, and other behaviors that are deliberately performed by actors as opposed to naturally occurring emotions experienced automatically by participants. However, there is little discussion available on the topic of the automatic or deliberate use of haptic behavior for emotion. In general, the use of touch is more regulated than other emotions (Argyle, 1988). For example, it may be extremely difficult for to prevent oneself from smiling during a funny movie, though not as difficult to prevent oneself from touching another human that one feels an affinity toward. In this sense, it could be the case that the use of touch to express emotion is more of a deliberate process than an automatic process. On the other hand, forcing oneself to touch someone for whom one has extremely negative behaviors may be extremely difficult—in this sense, using touch to deliberately override certain automatic emotions may be problematic. Although it is out of the scope of the work presented here to fully resolve this distinction, we focus on the use of touch to deliberately express emotions.

2. OVERVIEW OF EXPERIMENTS

The current set of studies attempts to understand how much emotion can possibly be transmitted from one person to another using a simple, force-feedback haptic device. Given that today's haptic devices are somewhat limited and it is quite difficult to create forces, surfaces, and dynamic movements similar to human touch, it is essential to investigate whether simple devices that are not analogs to human touch organs are capable of transmitting emotion.

In Experiment 1, 16 participants each generated seven different emotions (anger, disgust, joy, fear, interest, sadness, and surprise) by moving a two degree of freedom force-feedback joystick for 10 sec for each emotion. We analyzed the quantitative aspects of the force utilized in the different emotions.

In Experiment 2, 16 additional participants each interacted via VIT with the recordings of the previous participants and attempted to recognize each emotion. We analyzed the accuracy of the different emotions and compared human performance to a Support Vector Machine (SVM), a learning algorithm similar to a neural network that learned to automatically classify the seven emotions. The purpose of including the SVM was to determine how well an optimal classifier performed. By comparing the learning algorithm that is designed parse the emotions generated on any vector that separates the categories to human classification performance, we potentially gain insight concerning whether the any shortcomings in recognizing the emotions are due to generating the movements versus recognizing the movements.

In Experiment 3, 16 pairs of participants interacted in vivo, attempting to communicate the seven emotions to one another via a 10-sec handshake. We use this data as a baseline for the haptic interaction as well as to further understand the communication of emotion through touch.

It is important to note the current set of experiments is a preliminary exploration in the phenomenon of VIT. Our experiments have small sample sizes and the nature of the design is to be exploratory, more of a guide for directing future work as opposed to proving specific hypotheses.

3. EXPERIMENT 1: EMOTION GENERATION

3.1. Method

Design

In this study we sought to collect data on people's ability to represent various mental states and emotions using a force-feedback joystick. We manipulated a single variable within participants: emotion generated (anger, disgust, fear, interest, joy, sadness, surprise).

Materials and Apparatus

We used an Immersion Impulse Engine 2000 force-feedback joystick as the haptic device. The device provides movement along two degrees of freedom and is capable of outputting a maximum force of 2 lb (8.9 N). We placed the device on its side so that the handle faced toward the participant rather than toward the ceiling. One may call this position the "handshake position" because interacting with the device in this position is analogous to doing a handshaking motion. The joystick was secured to a table using clamps, and its height adjusted so that participants could interact with the joystick in a natural manner. Figure 2 shows the experimental setup.

Participants

Sixteen Stanford University undergraduates (9 male, 7 female) were paid for their participation in this study.

Procedure

In this study, we first acquainted participants with the joystick in a practice trial so that they had an idea of how interacting with the joystick would feel. For the practice trials, we instructed participants to interact with the joystick for 10 sec and then played back that same 10-sec recording to them. Participants then used the joystick for two more practice handshakes. The re-

Figure 2. **A user interacting with the Virtual Interpersonal Touch VIT device from this study.**

searcher then explained the real trials of the study. Participants were told that they would be expressing seven different emotions via the joystick. These seven emotions were based on Ekman and Friesen's (1975) work. Specifically, participants were asked to "do your best to communicate the exact mental state to someone else who may use the joystick to attempt to identify your specific mental states from the movement at a later date." We then began recording participants' attempts to convey emotions using the joystick. The order in which participants conveyed each of these emotions was randomized for each participant.

Participants were told that they would have up to 10 sec to express each emotion. For each trial, the researcher would verbally tell the participant the designated emotion. In all trials, a computer monitor counted up from 1 to 10 so that participants always knew how much time they had remaining to record. We allowed participants to record less than the entire 10 sec to avoid situations in which participants felt that they had conveyed the emotion well in 2 sec, for example, and then simply filled the rest of the 10 sec with motions that were not as focused on conveying the given mental state. We recorded data from the haptic joystick based on the *x–y* coordinates of the joystick every 5 msec during the trials. The joystick allowed participants to move freely, that is, it did not provide any resistance to their movements.

After each trial we asked participants to rate, on a scale from 1 (*extremely likely*) to 7 (*extremely unlikely*), how likely they felt that another person would be able to recognize the specific mental state or emotion they had just generated. We also asked participants to rate on a scale from 1 (*extremely well*) to 7 (*extremely poor*) how well they felt that they were able to express the given mental state or emotion using the joystick. At the end of the study, we also asked participants to write about what they thought about the task and whether they used any strategies to express the emotions via the haptic device.

Derived Measures

We computed a number of quantitative metrics from the recorded movements to analyze the data from the emotions. We describe each of these measures in turn.

- *Distance.* This metric is the total distance traversed by the tip of the joystick. A low score would mean that the participant barely moved the joystick, whereas a high score would mean that a lot of movement occurred.
- *Mean speed.* This metric is the average speed at which the participant moved the joystick. A low score would mean that the participant moved the joystick slowly, whereas a high score would mean that the participant moved the joystick very fast.
- *Standard deviation of speed.* This metric is the standard deviation of a participant's movement. A low score would mean a steady movement, whereas a high score would mean jerky movement.
- *Mean acceleration.* This metric is the average acceleration of a participant's movement. A low score would mean the participant was decelerating, whereas a high score would mean the participant was accelerating.
- *Standard deviation of acceleration.* This metric is the standard deviation of the acceleration of a participant's movement. The lower the score, the less change there was during the trial. The higher the score, the more the participant was speeding up and slowing down throughout the trial.
- *Angle.* This metric is the average angle of the major axis of the handshake from 0° to 180°. A score of 0° indicates a horizontal movement, 90° is straight up and down, and the angle moves counterclockwise as the score goes up.
- *Standard deviation of position.* This metric is the standard deviation of the joystick position on an x–y plane. A low score would mean staying close to a small area of the plane, whereas a high score would mean moving across many different areas of the plane.
- *Standard deviation of the major axis.* The major axis is the axis along which the average angle was made. The standard deviation of the major axis is

a measure of the deviation in position along the major axis. A low score would mean moving only very slightly along the major axis, whereas a high score would mean moving a great deal along the major axis.

- *Standard deviation of the minor axis.* The minor axis is the complement of the major axis. The standard deviation of the minor axis is a measure of the deviation in position along the minor axis. A low score would mean moving only very slightly along the minor axis, whereas a high score would mean moving a great deal along the minor axis.
- *Percent of major axis.* This metric is the ratio between the standard deviation of the major axis and the minor axis. A low score would mean comparable distances moved along both axes and thus an overall square or circular pattern. A high score would mean significantly more movement along one of the axes and thus an overall rectangular or oval pattern.

3.2. Results

Derived Measures

Figure 3 depicts plots of the seven emotions generated by the participants. We normalized scores ($M = 0$, $SD = 1$) on all of our derived measures before examining the differences between the seven emotions. To test whether different emotions produced significantly different scores on the derived measures, we ran a series of repeated measure analyses of variance (ANOVAs). Because we had seven emotions, a full post hoc pairwise comparison would have required us to calculate a family-wise error rate that took into account the fact that 21 comparisons were being made. Such a full comparison would require a test so conservative as to yield no significant pairwise comparisons in our small sample. In the following results, we list the overall significance for each repeated measure analyses and then describe the pairwise differences using a less conservative comparison of their 95% confidence intervals.

We ran a series of repeated measure ANOVAs using emotion as the independent factor and each of the derived measures as a dependent variable. The significance of each ANOVA is listed in Figure 4.

The 95% confidence interval plots for each derived measure are shown in Figure 5. With regards to distance, participants moved more when expressing Joy and Anger than most of the other emotions. On the other hand, participants moved noticeably less when expressing Sadness. The same pattern was repeated with average speed, standard deviation of speed, mean acceleration, and standard deviation of acceleration. Participants had a shorter major axis when expressing Fear than when expressing most other emotions. On the other hand, participants had a shorter minor axis when expressing Sadness and a longer minor axis when expressing Joy. Finally, participants had a more

Figure 3. **Plots of the 16 participants' movements for the seven emotions.** ***Note:*** **The outline around each box represents the limits of potential movements along the two dimensions. The maximum range in physical space for each dimension was approximately 28 cm.**

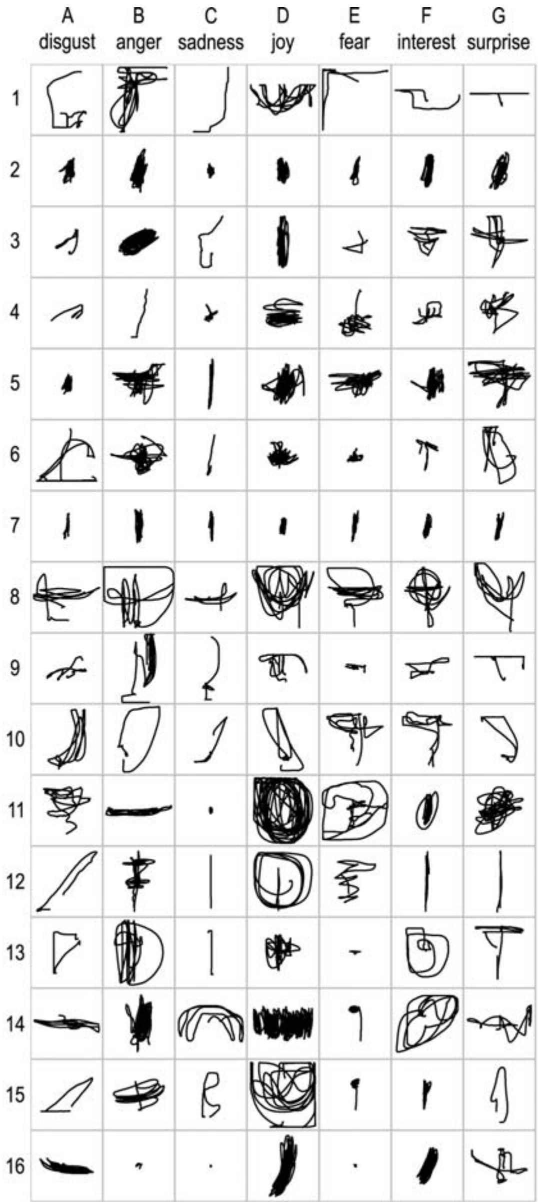
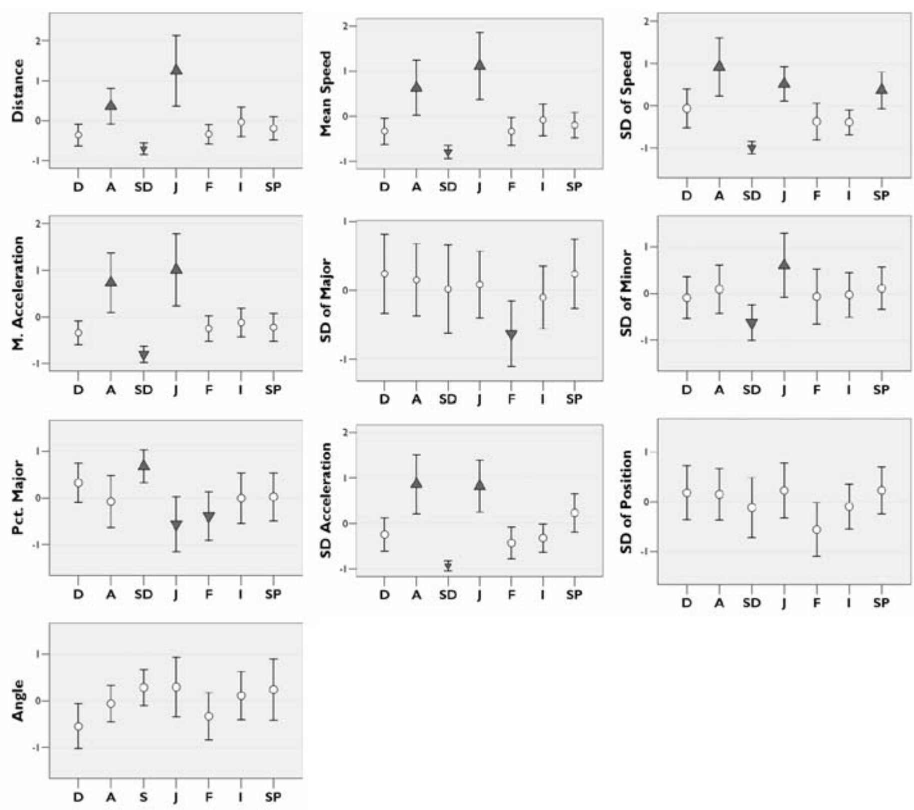


Figure 4. **Significance tests from repeated measure analyses of variance of all derived measures.**

Measure	F	p	η_p^2
Distance	11.78	$< .001$	.44
Speed (*M*)	13.10	$< .001$	.47
Speed (*SD*)	15.70	$< .001$	.51
Acceleration (*M*)	15.70	$< .001$	.45
Acceleration (*SD*)	15.68	$< .001$	.51
Angle	2.14	.06	.13
Position (*SD*)	2.11	.06	.13
Major axis (*SD*)	2.35	.04	.13
Minor axis (*SD*)	2.90	.01	.16
Major axis (%)	3.47	.004	.18

Figure 5. **The mean and 95% confidence intervals of the seven emotions across ten different metrics.** ***Note:*** **Bars denoted by solid arrows are significantly higher or lower than other bars.**



rectangular shape overall when expressing Sadness and more square shapes overall when expressing Joy and Fear. A summary of the differences is described in Figure 6.

Confidence Ratings

Participants were asked to give confidence ratings to each of their handshakes: "How difficult was it to generate this emotion?" and "How easily do you think someone else can recognize this emotion?" Because the average correlation between the two items was high (.79), we used their average as a composite confidence rating, with lower numbers indicating higher confidence. We performed an ANOVA to detect whether Emotion had an effect on the confidence ratings. The effect was not significant, $F(6, 90) = 1.76$, $p = .12$, $\eta_p^2 = .11$. The average confidence ratings are listed in Figure 7.

3.3. Discussion

The data from the emotion generation study suggest that there were indeed variances in handshaking behavior when different emotions were being expressed and that these variances can be quantified in meaningful ways. For example, sadness was expressed in slow, steady, and short movements, whereas joy was expressed in long, jerky, and fast movements. Given that different emotions were indeed expressed in measurably different ways, our next study explored how well participants could recognize emotions from these recorded handshakes as played back on haptic devices.

Figure 6. **Summary of differences in derived measures for the seven emotions.**

	Disgust	Anger	Sadness	Joy	Fear	Interest	Surprise
Distance	Short	Long	Short	Long	Short		
Speed (*M*)		Fast	Slow	Fast			
Speed (*SD*)		Jerky	Steady	Jerky			Jerky
Accel (*M*)		Faster	Slower	Faster			
Accel		High	Low	High			
Angle (*SD*)							
Position (*SD*)							
Major (*SD*)					Short		
Minor (*SD*)			Narrow	Wide			
Major (%)	Square		Rectangular	Square	Square		

Note. A label occurs for a given emotion on a measure when that emotion behaves in an extreme manner compared to the other emotions in terms of 95% confidence intervals. Accel = acceleration.

Figure 7. **Means and standard deviations of confidence scores across studies.**

	Exp. 1 Generation		Exp. 2 Detection		Exp. 3 Detection	
Emotion	***M***	***SD***	***M***	***SD***	***M***	***SD***
Disgust	4.84	1.99	4.75	1.61	2.94	1.52
Anger	3.89	1.44	4.50	2.10	3.35	2.32
Sadness	4.58	1.77	4.13	1.82	2.82	1.81
Joy	4.53	1.58	4.00	2.22	3.18	1.74
Fear	4.53	1.62	4.50	2.00	3.18	1.85
Interest	5.11	1.50	4.44	2.16	3.12	2.15
Surprise	5.00	1.27	4.56	2.03	3.53	1.70

Note. A label occurs for a given emotion on a measure when that emotion behaves in an extreme manner compared to the other emotions in terms of 95% confidence intervals. Exp. = experiment.

4. EXPERIMENT 2: HUMAN EMOTION RECOGNITION FROM HAPTIC DEVICES

4.1. Method

Design

In the second study, we sought to test people's ability to recognize others' attempts to convey mental states using the haptic joystick as well as their confidence in that recognition. The actual motions used to represent each mental state or emotion were those recorded in Experiment 1. Each participant received all of the recordings from a randomly selected previous participant from Experiment 1, such that all of the recordings were used exactly once.

Participants

Sixteen Stanford University undergraduates (9 male, 7 female) were paid for their participation in this study.

Procedure

We first acquainted participants with the joystick by playing them all the same recorded handshake motion on the joystick created as a template by the experimenters. This single handshake was not designated as any particular emotion; its purpose was simply to familiarize the participant with the haptic device by playing them a generic handshake. We then informed participants that they would be played a series of recordings and that, after each playing, they would be asked to try to identify which of the seven emotions was being conveyed. For each haptic recording they received, participants were not given any label at all. Instead, they were asked specifically to "do your best to determine the exact mental state that someone else was attempting to transmit." These seven emotions were listed on a sheet of paper and visible to the participant throughout the trials.

Participants received the same sequence of the seven emotions twice. In the first sequence, they were instructed to feel each recording and to think about what it may be without having to provide an answer. In the second sequence, after feeling each recording, participants were required to choose an emotion and indicate how confident they were in their choice on a scale from 1 (*extremely confident*) to 7 (*extremely unconfident*). Participants were allowed to respond with a given emotion only a single time. In other words, once they had used "joy" they were not allowed to use that emotion again. At the end of the trials, we asked participants to write about the task and what strategies they used to detect the mental states.

4.2. Results

On average, participants were correct on 33.04 % of trials. This was significantly above chance (14.29 %), $t(6) = 5.03$, $p < .002$. Figure 8 shows the responses and error rates by emotion. The percentage of hits and false alarms for each emotion is shown in Figure 9. We ran a one-way, repeated measures ANOVA with emotion as the independent factor and accuracy as the dependent variable. There was no significant difference between emotions, $F(6, 90) = .80$, $p < .58$, $\eta_p^2 = .05$.

We also examined participants' confidence ratings in their recognition of the emotions. We ran a one-way, repeated measures ANOVA with emotion as the independent factor and confidence ratings as the dependent variable. There was no significant difference between emotions, $F(6, 90) = .65$, $p < .69$, $\eta_p^2 = .04$. The average confidence ratings for each emotion is listed in Figure 7.

To test the ability to differentiate the generated emotions, we used a standard learning algorithm used to classify categorical data, the Torch3 SVM module with a radial basis function (RBF) kernel (see Doniger, Hofmall, & Yeh, 2002, for a similar use). We trained seven SVMs: one that separated "joy" from all other emotions, one that separated "sadness" from all other emotions, and so on. To classify a handshake, we then tested it against all seven SVMs and chose the best match. The parameters c and σ were tuned by dividing the participants into a 70% training group and 30% testing group and then using a gradient-ascent algorithm to determine which parameters trained the SVM to best match the test group. After attempting an exhaustive sampling of parameters, the best

Figure 8. **Average responses across 16 participants for the seven emotions.**

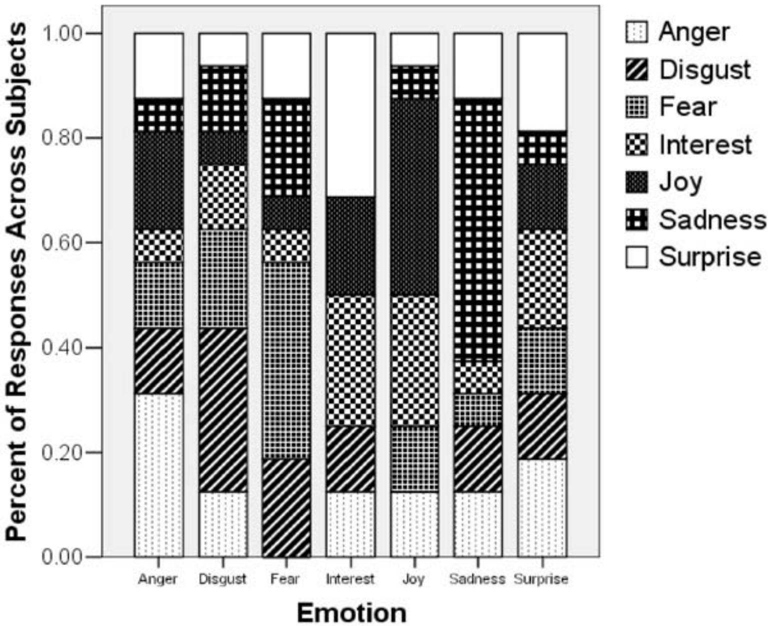


Figure 9. **Percentages of hits (percentage of responding with the correct emotion given the occurrence of the emotion) and false alarms (percentage of responding with the correct emotion given the nonoccurrence of the emotion) for each emotion.**

Emotion	Hits	False Alarm	Difference
Disgust	31.3%	11.5%	19.8%
Anger	31.3%	11.5%	19.8%
Sadness	50.0%	8.3%	41.7%
Joy	37.5%	10.4%	27.1%
Fear	37.5%	10.4%	27.1%
Interest	25.0%	12.5%	12.5%
Surprise	18.8%	13.5%	5.2%

results occurred near $c = 100$ and $\gamma = .00016$. The SVM was then tested by training on a random 70% of the participants and then classifying the remaining 30%. We repeated this train/test paradigm 1,000 times, choosing different combinations of participants each time. Across iterations, the SVM classified the handshake correctly 36.31% of the time, similarly to human participants.

4.3. Discussion

Our findings demonstrated that people were indeed able to recognize the emotions expressed via handshakes with accuracy approximately twice what would be expected by chance. It is interesting to note that for all of the seven emotions except for interest and surprise (which participants confused for one another), participants guessed the appropriate emotion more often than any of the incorrect alternatives. This pattern makes sense, as many of the frameworks of emotions do not include both interest and surprise as independent basic emotions (see Ortony & Turner, 1990, for a review).

In sum, participants were relatively astute at recognizing the emotions. In fact, a learning algorithm SVM designed purely to segment the seven emotions on any vector that separated the categories was only slightly more successful than the group of human participants. Consequently, it is most likely the case that the reason for error in detecting the emotions in Experiment 2 stemmed from the difficulty in generating an emotion through the haptic device in Experiment 1. In other words, given that the learning algorithm could not outperform a group of humans, there were most likely limited amounts of reliable information present in the digital motions that could be used to differentiate the emotions. Given the reduction of cues as compared with a real handshake—such as temperature and grip—we wanted to explore whether emotion recognition would improve when handshakes could be generated in a more naturalistic fashion with a wider range of emotional cues. Thus, in the third study, we repeated the study using two participants who shook hands in person.

5. EXPERIMENT 3: HUMAN EMOTION RECOGNITION FROM OTHER HUMANS

5.1. Method

Design

In the third study, we sought to test people's ability to recognize others' attempts to convey mental states through touch only while holding hands as well as their confidence in that recognition.

Participants

Thirty-two Stanford University undergraduates (16 male, 16 female) were paid for their participation in this study.

Procedure

Sixteen pairs of 2 participants engaged in a handshake through a doorway covered by a curtain, with each one on his or her own side of the curtain such that the only visual information received was about the hand and forearm. Each participant was monitored by a separate experimenter, who showed that participant specific instructions. To give the instructions silently to prevent the other participant from knowing the specific emotion, the experimenter pointed to areas of text while the participant read it to him- or herself.

We then informed one participant that he or she would be doing his or her best to generate emotions through the handshake. We informed the other participant that he or she would be asked to try to identify which of the seven mental states or emotions was trying to be conveyed. Participants received a randomized sequence of the seven emotions twice, each in the same random order. In the first sequence, they were instructed to generate and evaluate each emotion for the purpose of practice and to think about how to convey the emotions via touch. In the second sequence, after feeling each recording the participants performed the same confidence ratings as participants in Experiment 2. After the trials, the participants were asked to write about their subjective reactions to the task.

5.2. Results

Results demonstrated that people were quite good at recognizing the seven emotions through a handshake. The overall accuracy rate was 50.77%. This was significantly above chance (14.29 %), $t(6) = 14.42$, $p < .001$. We ran a one-way, repeated measures ANOVA with emotion as the independent factor and accuracy as the dependent variable. There was no significant difference between emotions, $F(6, 90) = .42$, $p < .86$, $\eta_p^2 = .02$. Figure 10 shows the responses and error rates by emotion. We next performed an ANOVA to detect whether Emotion had an effect on the confidence ratings of detection. The effect was not significant, $F(6, 90) = .38$, $p = .89$, $\eta^2 = .02$. The average confidence rating for each emotion is listed in Figure 7. The number of hits and false alarms for each emotion are listed in Figure 11.

6. GENERAL DISCUSSION

6.1. Summary of Results

In our study, we examined the ability of humans to transmit emotions via touch, both hand to hand and digitally mediated with VIT. When people

Figure 10. **Average responses across 16 pairs of participants for the seven emotions.**

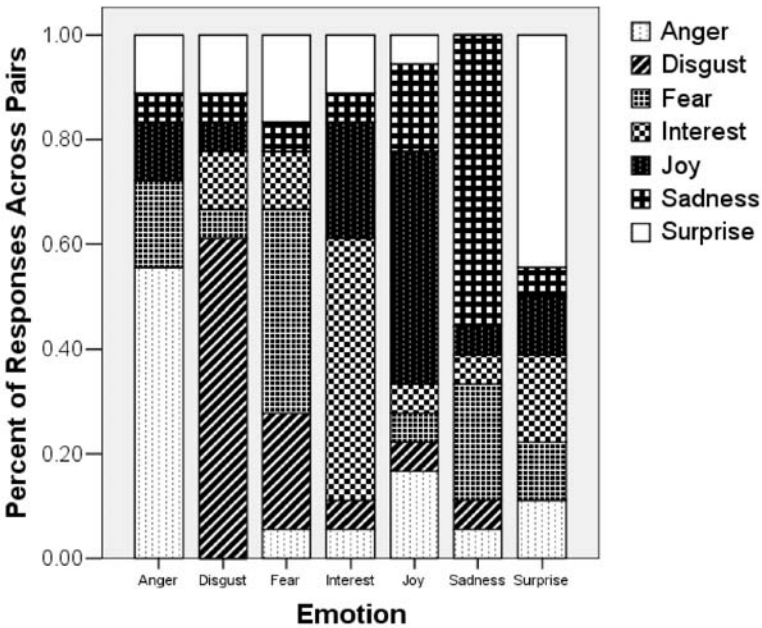


Figure 11. **Percentages of hits and false alarms for each emotion.**

Emotion	Hits	False Alarm	Difference
Disgust	55.6%	7.4%	48.1%
Anger	61.1%	6.5%	54.6%
Sadness	38.9%	10.2%	28.7%
Joy	50.0%	8.3%	41.7%
Fear	44.4%	8.3%	36.1%
Interest	55.6%	7.4%	48.1%
Surprise	44.4%	8.3%	36.1%

were asked to express emotions via haptic devices, we found reliable, predictable differences in measures that quantified different aspects of the movements (Experiment 1). Moreover, these differences were strong enough to allow other participants to interpret the emotional content of another person's handshake via haptic devices (Experiment 2) above chance. On the other hand, the reduction of physical cues in haptic devices as compared with in vivo handshakes—such as temperature and grip—lowered the accuracy of emotion detection (as seen in Experiment 3). Overall, our studies illustrate the ability of haptic devices to convey various emotions via a short interaction such as a handshake.

6.2. Limitations and Future Directions

There are a number of limitations to our work. First, we utilized only a simple two degree of freedom haptic device. In future work we plan to examine more elaborate devices that allow for more advanced transmission of the typical nonverbal cues and nuances that occur in interpersonal touch (Chaplin et al., 2000). Our choice of using a simple 2DOF device as opposed to a Phantom device that allowed more elegant movements was largely driven by the strategic goal of starting with an extremely simple device. However, using a more versatile haptic device to explore the generation and recognition of emotions is crucial for future work.

Furthermore, the task used in our study was not naturalistic. Instead of forcing people to generate deliberate emotions "on demand," just for the purpose of transmitting them, future studies should actually have participants experience the various emotions, for example, by having them watching emotional film clips (Rottenberg, Ray, & Gross, 2007) and then feature the transmission and reception of those actually experienced emotions via a haptic device. In other words, one criticism of our work is that we may not be studying actual, automatic emotions but only idealized, artificially constructed emotions. This distinction may be why our recognition rate (33%) was relatively low overall using the haptic device. However, these emotions generated "on demand" are still quite worthy of studying, as they may be the type that is utilized mostly during HCI (not unlike the use of standard emoticons in textual chat). Similarly, future work should also compare being able to generate novel emotional movements during VIT to choosing among preset movements designed to convey specific expressions.

One shortcoming of the work presented here is that it only examined VIT "in a vacuum." There is much reason to suspect that nonverbal and verbal behaviors are inextricably tied (Ekman, 1997; Kendon, 1970). It would be worthwhile to examine how the use of VIT to transmit emotions changes when accompanied by cues from other modes, such as voice, facial expressions, and other gestures.

Finally, there was some evidence that the current data support the idea of using a two-dimensional space to map emotions—hedonic valence and intensity. First, Table 1, which portrays how various emotions were characterized by specific movements, indicates that intensity and direction of movement were highly diagnostic of emotions, especially in differentiating sadness from other emotions (intensity) as well as differentiating anger from joy (horizontal movement). The anecdotal responses given by our participants on more than one occasion indicated that people used one or both of these dimensions in creating or recognizing the emotions, as can be seen in the appendix. For example, one participant noted, "Strategies I used included speed of up/down motion, length of hand shake and then force/strength." Consequently, future work should explore this mapping of emotions via haptic devices, perhaps by giving participants explicit instructions about strategies before generating the emotions.

6.3. Implications

This study has a number of implications for various aspects of HCI. First, it indicates that even with cues that are extremely degraded—for example, stripping a handshake of grip, temperature, dryness, texture, and other nonverbal cues—virtual interfaces can be effective at transmitting emotion. In other words, limited degree of freedom force-feedback devices can be used to add emotional content—such as warmth or intimacy—to virtual environments. This also opens to possibility of using haptic devices for the purpose of social influence such as, for example, using confident or cheerful handshakes to greet users. Also, previous research on emotion in user interface systems has shown that matching the mood of a car's warning voice to that of the mood of the car's driver (i.e., cheerful or sad) decreases the accident rate compared to when there is a mismatch (Nass & Brave, 2005). Perhaps haptic devices used to connect a user to an interface system can match the mood of users accordingly to enhance efficiency and productivity.

More important, the computer-mediated aspect of VIT means that touch communication can be altered and exploited in ways that are not possible in the real world. The ability of virtual environments in allowing nonveridical representation and interaction has been previously described as transformed social interaction (Bailenson et al., 2004). In nonmediated environments, we can shake hands with only one person at a time; using VIT, a user's handshake can be prerecorded and used for multiple greetings. Moreover, that handshake can be automatically tailored for the interactant. For example, one's handshake can be made firmer if another interactant prefers a firmer handshake. Given the social advantage that can be leveraged via mimicry (see Bailenson & Yee, 2005; Chartrand & Bargh, 1999), it would also make sense to strategically clone another person's handshake for future greetings with that person.

Our findings have demonstrated that humans touching one another virtually can transmit and receive emotional cues far above chance, and not too far off from what is possible in normal, face-to-face interaction. Furthermore, the data shed light on the development of haptic devices. By providing a quantitative framework for isolating aspects of different types of hand movements, our work assists other researchers in exploring the utility and theoretical possibilities of various types of virtual interfaces. Given the research in developing social robots for collaboration (Hinds, Roberts, & Jones, 2004) and education (Kanda, Hirano, Eaton, & Ishiguro, 2004), it is important to understand how haptic devices can be used to generate emotional content in other contexts as well, such as online gaming, training exercises, and chatrooms designed for the sole purpose of social interaction.

In sum, our work demonstrates that humans can express a range of emotions through hand-to-hand touch, whether that touch is computer mediated or not. Consequently, the use of VIT via haptic devices in all forms of computer-mediated communication should be strongly considered as the development of collaborative tools evolves. Given the power of touch in the physical world and the unique ability to amplify, multiply, and transform this power, it is inevitable that the theoretical underpinnings and applications of VIT receive attention.

NOTES

Acknowledgments. We thank Federico Barbagli, Ken Salisbury, and Hong Tan for helpful suggestions relevant to this research. Furthermore, we thank Keith Avila, Claire Carlson, Erin Dobratz, Alice Kim, Bryan Kelly, and Chelsea Maughan for their assistance with data collection.

Authors' Present Addresses. Jeremy Bailenson, Department of Communication, Stanford University, Stanford, CA 94305. E-mail: bailenson@stanford.edu. Nick Yee, Department of Communication, Stanford University, Stanford, CA 94305. E-mail: nyee@stanford.edu. Scott Brave, Baynote, Inc., 10051 Pasadena Avenue, Cupertino, CA 95014. E-mail: scott@scottbrave.com. Dan Merget, Department of Computer Science, Stanford University, Stanford, CA 94305. E-mail: merget@toast.net. David Koslow, Department of Computer Science, Stanford University, Stanford, CA 94305. E-mail: dkoslow@stanford.edu.

HCI Editorial Record. First manuscript received August 29, 2005. Revision received June 3, 2006. Accepted by Andrew Monk. Final manuscript received January 30, 2007. — Editor

REFERENCES

Argyle, M. (1988). *Bodily communication* (2nd ed.). London: Methuen.

Bailenson, J. N., Beall, A. C., Loomis, J., Blascovich, J., & Turk, M. (2004). Transformed social interaction: Decoupling representation from behavior and form in collaborative virtual environments. *PRESENCE: Teleoperators and Virtual Environments, 13,* 428–441.

Bailenson, J. N., & Yee, N. (2005). Digital chameleons: Automatic assimilation of nonverbal gestures in immersive virtual environments. *Psychological Science, 16,* 814–819.

Ball, G., & Breese, J. (2000). Emotion and personality in conversational agents. In J. Cassell, J. Sullivan, S. Prevost, & E. Churchill (Eds.), *Embodied conversational agents* (pp. 189–219). Cambridge, MA: The MIT Press.

Basdogan, C., Ho, C-H., Slater, M., & Srinavasan, M. A. (1998). The role of haptic communication in shared virtual environments. In J. K. Salisbury & M. A. Srinivasan (Eds.), *Proceedings of the Third PHANToM Users Group Workshop, PUG98* (AI Tech. Rep. No. 1643 and RLE Tech. Rep. No. 624) (pp. 443–460). Cambridge, MA: MIT Press.

Bradley, M. (2000). Emotion and motivation. In J. Cacioppo, L. Tassinary, & G. Brenston (Eds.), *Handbook of psychophysiology* (pp. 602–642). New York: Cambridge University Press.

Brave, S., Ishii, H., & Dahley, A. (1998). Tangible interfaces for remote collaboration and communication. *Proceedings of CSCW '98: Conference on Computer Supported Cooperative Work.* New York: ACM Press.

Brave, S., Nass, C., & Sirinian, E. (2001). Force-feedback in computer-mediated communication. *Proceedings of UAHCI'01: Universal Access in Human-Computer Interaction.* Mahwah, NJ: Lawrence Erlbaum Associates.

Brewster, S., & Brown, L. M. (2004). Tactons: Structured tactile messages for non-visual information display. *Proceedings of AUIC 2004.* Dunedin, New Zealand: Australian Computer Society.

Burgoon, J. K., & Walther, J. B. (1990). Nonverbal expectancies and the consequences of violations. *Human Communication Research, 17*, 232–265.

Cacioppo, J. T., Bernston, G. G., Klein, D. J., & Poehlmann, K. M. (1997). Psychophysiology of emotion across the life span. *Annual Review of Gerontology and Geriatrics, 17*, 27–74.

Cassell, J., & Thorisson, K. (1999). The power of a nod and a glance: Envelope vs. emotional feedback in animated conversational agents. *Applied Artificial Intelligence, 3,* 519–538.

Chang, A., O'Modhrain, S., Jacob, R., Gunther, E., & Ishii, H. (2002). ComTouch: Design of a vibrotactile communication device. *Design of Interactive Systems Conference.* London.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology, 76,* 893–910.

Chaplin, W. F., Phillips, J. B., Brown, J. D., Clanton, N. R., & Stein, J. L. (2000). Handshaking, gender, personality, and first impressions. *Journal of Personality and Social Psychology, 79,* 110–117.

Clynes, D. M. (1977). *Sentics: The touch of the emotions.* New York: Anchor Press/Doubleday.

Collier, G. (1985). *Emotional expression.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Crusco, A. H., & Wetzel, C. G. (1984). The Midas touch: The effects of interpersonal touch on restaurant tipping. *Personality and Social Psychology Bulletin, 10,* 512–517.

Doniger, S., Hofmall, T., & Yeh, J. (2002). Predicting CNS permeability of drug molecules: Comparison of neural network and support vector machine algorithms. *Journal of Computational Biology, 9,* 849–864.

Durlach, N., & Slater, M. (2000). Presence in shared virtual environments and virtual togetherness. *Presence: Teleoperators and Virtual Environments, 9,* 214–217.

Ekman, P. (1997). Should we call it expression or communication? *Innovations in Social Science Research, 10,* 333–344.

Ekman, P. (2001). *Telling lies: Clues to deceit in the marketplace, politics, and marriage.* New York: Norton

Ekman, P., Davidson, R. J., & Friesen, W. V. (1990). The Duchenne smile: Emotional expression and brain physiology II. *Journal of Personality and Social Psychology, 58,* 342–353.

Ekman, P., & Friesen, W. V. (1975). *Unmasking the face. A guide to recognizing emotions from facial clues.* Englewood Cliffs, NJ: Prentice-Hall.

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

Ekman, P., & Friesen, W. V. (1978). *Facial action coding system: A technique for the measurement of facial movement.* Palo Alto, CA: Consulting Psychologists Press.

Essa, I., & Pentland, A. (1994). A vision system for observing and extracting facial action parameters. *Proceedings of IEEE Computer Vision Pattern Recognition Conference 1994.* Seattle, WA: IEEE.

Fogg, B. J., Cutler, L., Arnold, P., & Eisback, C. (1998). HandJive: A device for interpersonal haptic entertainment. *Proceedings of CHI '98: Conference on Human Factors in Computing Systems*, New York: ACM.

Fridlund, A. (1994). *Human facial expression: An evolutionary view.* San Diego, CA: Academic.

Fussell, S., Setlock, L., Yang, J., Ou, J., Mauer, E., & Kramer, A. (2004). Gestures over video streams to support remote collaboration on physical tasks. *Human-Computer Interaction, 19*, 273–309.

Goldberg, K., & Wallace, R. (1993). Denta-Dentata. *Visual Proceedings of SIGGRAPH '93: International Conference on Computer Graphics and Interactive Techniques*. New York: ACM.

Hansson, R., & Skog, T. (2001). The LoveBomb: Encouraging the communication of emotions in public spaces. *Computer-Human Interaction (CHI) 2001 Extended Abstracts*. Seattle, WA: ACM Press.

Hinds, P., Roberts, T., & Jones, H. (2004). Whose job is it anyway? A study of human-robot interaction in a collaborative task. *Human-Computer Interaction*, 19, 151–181.

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

Kanda, T., Hirano, T., Eaton, D., & Ishiguro, H. (2004). Interactive robots as social partners and peer tutors for children: A field trial. *Human-Computer Interaction*, *19,* 61–84.

Kendon, A. (1970). Movement Coordination in Social interactions. *Ada Psychologica, 32,* 101–125.

Kim, J., Kim, H., Tay, B., Manivannan, M., Srinivasan, M., Jordan, J., et al. (2004). Transatlantic touch: A study of haptic collaboration over long distance. *Presence: Teleoperators and Virtual Environments, 13,* 328–337.

Lanier, J. (2001, April). Virtually there. *Scientific American,* pp. 66–75.

Mehrabian, A. (1967). Orientation behaviors and nonverbal attitude communication. *Journal of Communication, 17,* 324–332.

Michel, P., & El Kaliouby, R. (2003). Real time facial expression recognition in video using support vector machines. *Proceedings of the 5th International Conference on Multimodal Interfaces (ICMI)*. New York: ACM.

Murray, I. R., & Arnott, J. L. (1993). Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *Journal of the Acoustical Society of America, 93,* 1097–1108.

Nass, C., & Brave, S. (2005). *Wired for speech: How voice activates and advances the human-computer relationship.* Cambridge, MA: MIT Press.

Noma, H., & Miyasato, T. (1997). Haptic communication for cooperative object manipulation. *Proceedings of the International Workshop on New Media Technology.*

Oakley, I., Brewster, S. A., & Gray, P.D. (2000). Communicating with feeling. *Proceedings of the First Workshop on Haptic Human-Computer Interaction.* London: Springer Verlag.

Ortony, A., & Turner, T. J. (1990). What's basic about basic emotions? *Psychological Review, 97,* 315–331.

Pantic, M., & Rothkrantz, L. J. M. (2003). Toward an affect-sensitive multimodal human-computer interaction. *Proceedings of the IEEE.*

Picard, R. W. (1997). *Affective computing.* Cambridge, MA: The MIT Press.

Picard, R. W., & Bryant Daily, S. (2005). Evaluating affective interactions: Alternatives to asking what users feel. *CHI Workshop on Evaluating Affective Interfaces: Innovative Approaches.* New York: ACM.

Plutchik, R. (2001). The nature of emotions. *American Scientist, 89,* 344.

Richardson, D. C., Spivey, M. J., Edelman, S., & Naples, A. D. (2001). "Language is spatial": Experimental evidence for image schemas of concrete and abstract verbs. *Proceedings of the Twenty-third Annual Meeting of the Cognitive Science Society.* Mahwah, NJ: Lawrence Erlbaum Associates.

Rottenberg, J., Ray, R. R., & Gross, J. J. (2007). Emotion elicitation using films. In J. A. Coan & J. J. B. Allen (Eds.), *The handbook of emotion elicitation and assessment* (pp. 87–108). London: Oxford University Press.

Rovers, A. F., & Essen, H. A. v. (2004). HIM: A framework for haptic instant messaging. *Proceedings of ACM SIGCHI (CHI '04).* New York: ACM.

Sallnas, E., Rassmus-Grohn, K., & Sjostrom, C. (2000). Supporting presence in collaborative environments by haptic force feedback. *ACM Transactions on Computer-Human Interaction (TOCHI).* New York: ACM.

Scherer, K. R. (1989). Vocal measurement of emotion. In R. Plutchik & H. Kellerman (Eds.), *Emotion: Theory, research, and experience* (Vol. 4, pp. 233–259). San Diego: Academic.

Sproull, L., & Kiesler, S. (1986). Reducing social context cues: Electronic mail in organizational communication. *Management Science, 32,* 1492–1512.

Strong, R., & Gaver, B. (1996). Feather, scent and shaker: Supporting simple intimacy. *Videos, Demonstrations, and Short Papers of CSCW'96: Conference on Computer Supported Cooperative Work.* New York: ACM.

Turk, M., & Kölsch, M. (2004). Perceptual interfaces. In G. Medioni & S. B. Kang (Eds.), *Emerging topics in computer vision* (pp. 455–519). Upper Saddle River, NJ: Prentice Hall.

Watson, J. (1930). *Behaviorism* (2nd ed.). New York: Norton.

White, N., & Back, D. (1986). *Telephonic arm wrestling*. Shown at the Strategic Arts Initiative Symposium, Salerno, Italy.

Whittaker, S. (2002). Theories and methods in mediated communication. In A. Graesser, M. Gernsbacher, & S. Goldman (Ed.), *The handbook of discourse processes* (pp. 243–286). Mahwah, NJ: Lawrence Erlbaum Associates.

APPENDIX. SELECTED ANECDOTAL RESPONSES

A1. Selected Anecdotal Responses From Experiment 1

- Strategies I used included speed of up/down motion, length of hand shake, and then force/strength.
- A lot was subject to interpretation though, since I don't usually move my hands to convey mental states.
- I used strategies like trying to feel my assigned emotion in order to convey my mental state.
- Some were particularly difficult to distinguish (i.e., anger vs. disgust)—I found myself making the actual emotion faces while moving my hand in order to make the task easier.
- The hardest part was thinking of the context for the device and it made it hard to convey emotion to it cause it didn't have any convincing physical presence to me.
- It was hard to depict the distinction between different mental states because the handshake machine gave no resistance and also because it was unable to record the grip or firmness of how tight I was holding the "hand."

A2. Selected Anecdotal Responses From Experiment 2

- Not having a sense for group attained by clasping fingers made it difficult to be entirely sure of an emotion.
- Are short, sharp motions angry or surprised or what?
- Most mental states aren't expressed through hand movements, so it was difficult to ascribe arbitrary motions to specific emotions.
- It was easy to notice intensity on each shake. It was hard to imagine feeling without facial expression or language.