Concurrence of Facial and Bodily Expression: A Feasibility Study

Elizabeth Crane
Division of Kinesiology
University of Michigan
Ann Arbor, MI 48109 USA
734 763 0013
bcrane@umich.edu

Melissa Gross
Division of Kinesiology
University of Michigan
Ann Arbor, MI 48109 USA
734 763 0013
mgross@umich.edu

Barbara Fredrickson
Department of Psychology
University of Michigan
Ann Arbor, MI 48109 USA
734 936 0640
blf@umich.edu

ABSTRACT
This paper discusses the challenges of studying affective interactions and reviews a pilot study designed to test the feasibility of studying the concurrence of facial and bodily expression. The study had three goals. First, evaluate the use of our emotion elicitation method. Second, determine if a specialized head mount for a face camera worn by participants interfered with performance. Third, determine if facial expression occurred without social interaction. The emotion elicitation method was successful for four target emotions (anger, sad, content, and joy). The specialized head mount did not interfere with emotion elicitation and the subjects displayed evidence of facial expression.

Keywords
Facial expression, multimodal communication, affective interactions

INTRODUCTION
Applications that seamlessly interact with humans require an intelligent system that identifies and "understands" subtle cues about a person's internal state. A growing body of research exists describing and interpreting cues from individual affective channels [1] [2] [3] [4] [5] [8] [15] [16]. Research on facial expression is one example [1] [2] [3] [4]. Knowledge gained from studying facial expression has influenced fields beyond psychology. For example, law enforcement agencies are interested in subtle cues detected in facial expression indicating if a person is lying [6] [7] [11]. The entertainment industry is interested in understanding facial expression to create better animations [14]. Product developers study facial expressions to help determine if people like the product or not. Each application of facial expression measurement or creation requires technology to complete the task. The ability to detect and interpret both obvious and subtle expressive cues is becoming increasingly intertwined with technology. Developing technology with some level of expressive knowledge is a logical progression given that humans are expressive by nature and that we communicate with more than just words.

In human – human interactions, however, we generally have access to more than facial expression to help detect an affective state. Knowledge of a single channel does not account for or describe interactions with additional channels such as bodily expression, voice, or other physiological cues. Therefore, research is needed to understand both individual channels and how channels interact [12]. By doing this we will develop a more complete understanding of the subtle cues indicating a person's affective state. One factor hindering the development of intelligent affective systems is the uncertainty around how multiple channels work together. This paper discusses the challenges of studying affective interactions and reviews a pilot study designed to test the feasibility of studying the concurrence of facial and bodily expression.

BACKGROUND
One difficulty of working with and studying affective interactions is defining a framework for understanding affective states. Larsen and Fredrickson [9] advocate for the use of a working definition. With a working definition we can construct a framework to ask experimental questions, design methodologies, and interpret results.

Our working definition includes four assumptions. First an affective state is a multifaceted process [10]. One part of this process is an occurrence of an internal or external event. An example of an external event is interacting with a computer that is not functioning properly. An example of an internal event is thinking about previous or future experiences but not directly interacting with the surrounding environment. The second assumption is that an event (internal or external) can cause a specific internal affective state. The internal state is not directly measurable [9]. Third, the affective state can cause a measurable output in one or more affective channels [5] [9] [15] [16]. However, the measurable output is not the internal state. The measurable output is a consequence of the internal state. In other words, there is a gap between what sensors detect and the actual affective state. The computer system must use the measurable outputs to make an inference about the internal state of the person. This is a step removed from the actual internal state. Finally, the fourth assumption is that internal states are complex and we assume that there may be more internal states than there are ways to express these states (or at least accurately differentiate the states).

Thus, the only parts of this process we can evaluate in HCI research are the measurable responses (i.e., body movements, voice tone, facial expressions) but not the actual internal state [1] [2] [5] [15] [16]. Therefore, understanding how affective states relate to this process may allow us to select appropriate sensors to detect an affective state and possibly allow computers to elicit affective states. If the wrong affective channel or combination of channels is used for detecting a specific state the system may fail (i.e., the human – computer interaction may be unsuccessful).

A second challenge is to determine how much and what type of affective information is necessary to accurately communicate or recognize an internal state. For example, does more information increase recognition rate? To begin answering this and related questions we need to study the concurrence of affective channels.

In a pilot study we tested the feasibility of studying the concurrence of facial and bodily expression. The study addressed three questions. First, does the autobiographical memories paradigm successfully elicit our target emotions in a laboratory setting? Second, does wearing a head mount with a small camera attached to capture facial expression interfere with emotion induction and/or facial expression? Finally, do people make facial expressions in a laboratory setting in the absence of social interaction during the task?

METHODS

24 university students (15 female, 9 male; 18-28 yrs) volunteered to participate after giving informed consent. An autobiographical memories paradigm was used to elicit five target emotions (anger, sad, neutral, content, or joy). Each participant was asked to write down events from his/her own life in which he/she felt each of four target emotions, and no emotion at all (neutral). Prior to each movement trial, the participant was asked to recall one of these experiences. No instructions were given about the speed of the movement. Participants wore a special motion capture suit (tight-fitting dark clothes) with 46 reflective markers that demarcated head, neck, shoulder, arm, forearm, hand, torso, pelvis, thigh, leg and foot segments.

After recalling a target emotion, participants walked across the lab (approximately three meters). Motion data were acquired using a video-based, 6-camera motion capture system (EvaRT, Motion Analysis Corp.) at a sample rate of 60 Hz. Front and side view video was recorded for all participants. Six of those participants also wore a special camera designed to record video of facial expression (Adventure Cam II, Viosport). Participants performed three trials for each emotion in a block and their self-selected best trial was included in the analysis.

Participants rated the intensity of their "felt" emotions after each trial using a questionnaire. The feelings questionnaire included the four target emotions (e.g., anger - "I felt angry, irritated, annoyed.") and 4 non-target "distracter" emotions (e.g., fear - "I felt scared, fearful, afraid."). A 5-item Likert scale (0 = not at all; 1 = a little bit; 2 = moderately; 3 = a great deal; 4 = extremely) was used to score intensity. Intensity scores of two ("moderately") or greater were considered a "hit" (the subject felt the emotion). Neutral trails were considered neutral if all 8 scores were less than two.

Facial Action Coding System (FACS) was used to code the occurrence of facial expression. Because subtle changes were expected the full FACS coding was used.

RESULTS

The results reported here are preliminary. A more complete analysis of the data collected is in progress.

Self report of felt emotion

Median scores were calculated for three groups: all subjects, subjects that did not wear the head mount, and subjects that wore the head mount. The subject's self-selected best trial was used to calculate the median scores. For each group the median score for anger, sad, joy, and content was four. Out of 96 trials (four emotions x 24 subjects) only one subject reported a score less than two (score = 1, emotion = content). Neutral trials, however, were not always neutral. Thirteen of the 24 (54.2%) subjects' self-selected best neutral trials were neutral. In the group that did not wear the head mount 9 of the 18 (50%) trials were neutral compared to 4 of the 6 (66.7%) trials in the group that wore the head mount.

Occurrence of facial expression

FACS coding was used to evaluate each of the faces for evidence of emotion. With the exception of one trial all neutral trials were neutral. One trial showed some evidence of a smile.

All trials of joy and content displayed evidence of emotion. A smile was present in both joy and content. However, the intensity of the smile was greater in joy. Four of the six joy trials also displayed a cheek raise and eye lid compression. Eye movement was coded differently for joy and content. In joy the eyes tended to remain focused on a point in front of the walker. In content, the eyes tended to look around.

Four of the six walkers displayed facial evidence for sad. In these four walkers there is evidence of a slight jaw drop with the lips closed, eye looked down, and eye lids closed so that the iris was no longer visible.

None of the anger trials showed the traditional expression for anger, i.e. lowering the brow and tightening the eye lids. All, trials except one looked neutral. The one trial that showed some expression was a tightening of the lips that was not related to swallowing.

CONCLUSIONS

The criteria for a "felt" emotion was met or exceeded in all trials studied except for one. This suggests that the autobiographical memories paradigm successfully elicited the target emotions in our walkers. The least successful emotion was no emotion or neutral. Further analysis of the self-report questionnaires needs to be done to determine which emotion(s) were present in the neutral trials. Because of the importance of having a neutral trial available for comparison, we would like a greater success rate in our neutral trials.

The similarity of the median self-report scores between the two groups (head mount and no head mount) suggests that the head mount did not interfere with emotion elicitation.

This preliminary analysis also suggests that facial expression does occur in the laboratory setting without social interaction during the walking trial. The subtle changes and often low-intensity facial expression suggest full FACS coding is necessary.

DISCUSSION

Affective interaction and technology are becoming increasingly intertwined. We need technology to help understand affective states and we need to understand affective states to advance certain technologies. Both objectives require a framework for thinking about and working with affective states. Because more than one channel of affective information is often available we need to understand them individually and their concurrent interactions with additional channels. This feasibility study suggests methods that work in a laboratory setting. However, we may not be able to treat all affective states equally. One problem that occurs when studying more than one channel is the increased technology required to detect physiological and behavioral changes. We must be sure that the technology has minimal interference with the natural behavior.

Further analysis is also underway to study observer recognition of walker affect from front and side views. In addition we are studying the walker kinematics to determine subtle expressive behavior in body movement.

ACKNOWLEDGMENTS

The authors thank everyone that helped with data collection Melanie Cluss, Zara Schulman, Jon Dollison, and Zach Webster, and Alex Flick. A special thanks to those that wore the special motion capture suit during protocol testing.

REFERENCES

1. Ambador, Z., Schooler, J. W., Cohn, J. (2005) Deciphering the enigmatic face: The importance of facial dynamics in interpreting subtle facial expressions. *Psychological Science*, 16(5). 403-410.
2. Atkinson, A.P., Winand, H.D., Gemmell, A.J., Young, A.W. (2004) Emotion perception from dynamic and static body expressions in point light and full light displays. *Perception*, 33(6). 717-746.
3. Bassili, J.N. (1979) Emotion recognition: The role of facial movement and the relative importance of upper and lower areas of the face. *J. Personality and Social Psychology*, 37 (11). 2049-2058.
4. Crane E., Gross M., Fredrickson B. L., Koditschek D., Gerstner G., Validation of Emotion in Body Movements. *International Society for Research on Emotions*. (New York, NY, 2004).
5. Ekman, P. E. (1989). Arguments and evidence for universals in facial expressions of emotion. In Wagner, H. and Manstead, A. (eds) *Handbook of Social Psychophysiology.* Wiley & Sons, New York.
6. Frank, M. Ekman, P. (2004) Appearing Truthful Generalizes Across Different Deception Situations. *Journal of Personality and Social Psychology,* 86(3). 486-495.
7. Frank, M. Ekman, P. (1997) The Ability to Detect Deceit Generalizes Across Different Types of High-Stake Lies. *Journal of Personality and Social Psychology,* 72(6). 1429-1439.
8. Gross, M. Fredrickson, B.L., Koditschek, D.E., Gerstner, G.E. (2004) Kinematics of emotion in body movements. *International Society for Research on Emotions*. (New York, NY, 2004).
9. Larsen, R. J. & Fredrickson, B. L. (1999). Measurement issues in emotion research. In Kahneman, D., Diener, E. and Schwarz, N. (eds) *Well-being: Foundations of hedonic psychology*. Russell Sage, New York.
10. Levenson, R.W. (1994). Human emotion: A functional view. In Ekman, P., Davidson, R.J. (eds) *The nature of Emotion: Fundamental questions.* Oxford University Press, New York.
11. Navarro, J. (2003). A four-domain model for detecting deception: an alternative paradigm for interviewing. *The FBI Law Enforcement Bulletin*, 72(6). 20-26.
12. Pantic, M. Sebe, N. Cohn, J. Huang, T. Brave new topics 2: affective multimodal human-computer interaction: Affective multimodal human-computer interaction. *13th ACM international conference on Multimedia MM '0.* (Hilton, Singapore, 2005). 669-676.
13. Picard, R.W. (1997). Affective computing. M.I.T. Press, Cambridge MA.

14. Pyun, H. Kim, Y. Chae, W. Kang, H. Shin, S. Facial animation & hair: An example-based approach for facial expression cloning. *ACM SIGGRAPH / Eurographics symposium on Computer animation SCA '03*. (San Diego, California, 2003).

15. Scherer K.R. (2003) Vocal communication of emotion: A review of research paradigms. *Speech Communication*, 40(1-2). 227-256.

16. Wallbott HG (1998) Bodily expression of emotion. *European Journal of Social Psychology,* 28(6). 879-896.