Implementation of FACS for synthetic characters for use in studying facial expression recognition by survivors of childhood cancer

Robert C. Hubal (1), Ph.D., Noah R. Evens (1), David P. FitzGerald (2), Ph.D., Kristina K. Hardy (2), Ph.D., Victoria W. Willard (2), & Melanie J. Bonner (2), Ph.D.

(1) RTI International, Research Triangle Park, NC, USA
(2) Duke University Medical Center, Durham, NC, USA

Corresponding author:

Robert Hubal
RTI International
3040 Cornwallis Road
Research Triangle Park, NC 27709-2194
919/541-6045
919/541-6965 (fax)
rhubal@rti.org

ABSTRACT

A significant portion of pediatric cancer survivors will experience cognitive, academic, and social difficulties that limit their quality of survival well into adulthood. Of these, the least is known about the nature and extent of survivors' social problems. What is known about survivors' difficulties with nonverbal processing implies that their social lives may be hindered by problems with nonverbal aspects of social communication (e.g., interpreting facial expression, body language, tone of voice). Because of limitations inherent to current methodology and measurement, however, important information about the link between survivor's nonverbal deficits and their social functioning is unknown. As part of a set of studies to assess deficits associated with poor social functioning in childhood cancer survivors as compared to healthy children, we are developing a new instrument involving a facial expression recognition task. Our instrument employs Facial Action Coding System action units to systematically manipulate facial expressions, and an easy-to-use interface for the target pediatric population.

BACKGROUND AND SIGNIFICANCE

Despite increasingly favorable prognoses for survival of childhood cancer, these patients are at high risk for both acute and late-occurring sequelae associated with their disease and treatments. While all long-term pediatric patients are at high-risk for both late-occurring physiologic and psychosocial effects of cancer therapy, children who receive therapies that impact the central nervous system are at even higher risk for cognitive, social, and psychological disorders (Moore, 2005). Various studies have shown between 40 and 100 percent of survivors of pediatric brain tumors will evidence some sort of cognitive deficit resulting from a combination of disease and treatment variables (see Mulhern & Palmer, 2003, for a review). In addition, at least 30 percent of survivors of blood-based cancers will experience some degree of neurocognitive deficit (Copeland, et al., 1996). The repercussions of these deficits can be lasting and costly, with many childhood cancer survivors never achieving the normal milestones of adulthood, such as living independently, marrying, and procuring stable employment (Maddrey, et al., 2005; Zebrack, et al., 2002, 2004). Hence, some investigators have called for better assessment of critical psychosocial variables associated with a survivor's ability to successfully integrate into society.

Effective social interaction requires focused attention to and interpretation of complex and varied nonverbal social cues including facial expressions, body language, and tone of voice. The largest part of social communication is nonverbal (Knapp, 1972; Mehrabian, 1971), and in particular, facial expressions are one of the richest sources of nonverbal social information (Blair, 2003). Other models (e.g., Crick & Dodge, 1994; Lemerise & Arsenio, 2000) posit that decoding of facial expression represents the first two steps to accurately understand and react to a social situation, and that errors at these levels have potentially negative repercussions for social interactions (e.g., misinterpreting a smile of polite attention as one of genuine interest). Given that survivors often have nonverbal cognitive deficits (Buono, et al., 1998; Carey, et al., 2001), it is reasonable to assume that they may make more of these types of social errors.

In preliminary studies we used the Diagnostic Analysis of Nonverbal Accuracy – 2 (DANVA2; Nowicki, & Duke, 2001) to assess facial expression recognition in pediatric survivors of brain tumors versus children with Juvenile Rheumatoid Arthritis, and found robustly significant differences between the groups (Bonner, et al., in press). The DANVA2 is a facial recognition task consisting of 48 photographs of adult and child faces, depicting four basic expressions of happiness, sadness, anger, and fear. For each expression, both low- and high-intensity faces are included. Participants are shown each photograph and asked to identify the expression depicted as quickly as possible. While the DANVA2 has shown empirical efficacy for our initial understanding of the apparent complex nonverbal skill deficits in this population, it is an inherently unstandardizable measure that may only tell part of the story. Because the DANVA2 involves actors who were asked to display the appropriate expression after reading a vignette, it is not possible to ensure that the validity or intensity of the facial expressions is standard across photos. Moreover, it is not possible to dynamically adjust the characteristics or expressions of the DANVA2 faces. Thus, errors made by cancer survivors could reflect the actors' failure to accurately display the required expression, particularly for low-intensity emotions which may be harder to represent (especially for child actors).

FACIAL EMOTION RECOGNITION TOOLS

Recent technological advances have led to the emergence of more sophisticated and innovative methods for assessing children's processing of facial expressions. For example, a study of the effects of maltreatment on children's ability to perceive anger (Pollack & Sinha, 2002) used a computerized method that presents an image starting with an undifferentiated face that gradually gains organization and resolution to form a coherent facial expression. With this innovative technology, researchers found that children with a history of maltreatment recognized anger expressions with less visible information than other children, suggesting that they are sensitized and vigilant to threat cues. Blair and colleagues (2001; Coupland, et al., 2004) have devised the Emotional Expression Multimorph Task that uses images of nine models, each portraying six basic emotions, with expressions manipulated to 40 increments between a neutral state and the full expression. Participants respond as soon as they identify the target emotion during the morph. The tool was capable of yielding deficits in the recognition of facial sadness and disgust in children with psychopathic tendencies, and in establishing links between affect and thresholds for recognizing happy or disgusted faces. Massaro conducted studies using variations of features of a face to assess how children and adults distinguish one face from another (Schwarzer & Massaro, 2001; see also George, Scaife, & Hole, 2000). Massaro has also used a cartoonish face called Baldi within a number of studies involving children (see, e.g., Massaro & Bosseler, 2006). A program called Let's Face It! (Tanaka, et al., 2003) was designed for developing autistic children's face processing skills and normalizing neurological face recognition functions, using pictures of faces and facial components that the child manipulates in several tasks, including recognizing facial expressions and assigning labels to various facial expressions.

However, none of these tools enables the systematic manipulation of an unlimited number of facial expressions, which we feel is important for our research and others' involving social skill assessment. For instance, we were surprised to find no tool using the intensively studied action units that underlie the Facial Action Coding System (FACS; Ekman, Friesen, & Hager, 2002), nor were we able to obtain a tool using non-cartoonish faces that manipulated facial expressions using action units. Thus, we set out to create this tool. The remainder of this paper details the development and addresses system- and user-testing issues surrounding the use of our FER instrument.

FACIAL EXPRESSION RECOGNITION INSTRUMENT (FERI): Rationale and Features

We are systematically developing facial expressions by referring to FACS criteria. FACS uses the movement of facial muscle groups (action units; AU's) to measure facial expression. The absence of expert tools that aid in defining the AU's associated with FACS led to an investigation of a similar encoding scheme, the Moving Pictures Experts Group Facial Animation standard (MPEG-4 FA). Despite the existence of capable tools supporting MPEG-4 FA, their use was limited by subtle differences between the musculature coding of FACS AU's and MPEG-4 Facial Animation Parameters (FAP's). As we were disinclined to devise a complicated mapping between AU's and FAP's, particularly for the child faces that we needed, we followed a more traditional approach of defining AU's using a variety of mesh deformation techniques, rendered as a series of animation keyframes.

The client application is a derivative of Xface (http://xface.itc.it/), an open source project based on the MPEG-4 standard. Modifications to the Xface code and user interface enabled real-time manipulation and blending of dozens of AU's, and caching of the composite of their weighted sum. The initial 3D head known as Alice, the reference model delivered with Xface, was produced using Singular Inversions' FaceGen, the head and face mesh generation application that was already an integral tool in our research product line.

Though FaceGen will output visemes, phonemes, and expressions that we use in other applications (see, e.g., Hubal, Kizakevich, & Furberg, 2007), for this application these files went unused, and instead the head mesh was imported into Autodesk 3ds Max for further manipulation. For each AU, a corresponding vertex selection set was identified and named for later retrieval, though certain AU's representing incongruent muscle groups were composed of the aggregation of multiple selections. AU's mirrored along the vertical axis were divided into independent left and right selections. With the rest mesh at frame zero, animation keyframes were captured, each representing the extreme position of one AU. These deformations were applied to a control mesh, a copy of the original Alice mesh that is mapped against the final 'presentation' mesh, allowing multiple faces to be exported without having to recreate the AU's for each. The Max files are exported as a VRML Flip-book, a native import format of Xface, resulting in one VRML file for every AU.

We chose a subset of AU's to implement, specifically those that are most implicated in facial expression. For example there are a number of AU's associated with the eyes and eyebrows (inner- and outer-eyebrow raising, eyebrow lowering, upper eyelid raising, cheek raising and eyelid compressing, eyelid tightening, all of these for both left and right) that independently or collectively help define different facial expressions. The same is true for AU's around the nose and mouth. The FACS manual describes certain combinations of AU's when those combinations are not necessarily able to be linearly added; for the moment we enable these certain combinations by allowing for the computation of a weighted composite sum.

We are developing an interface surrounding Xface to enable any application designer to define and label those weighted combinations of AU's (actually, the keyframes representing each AU). This interface is important because we do not want our participants to be able to manipulate the AU's individually. Instead, as in some other products, we wish to present to the participant with a single slider bar. The movement of this bar would cause a change between a state signifying a neutral expression and a state (e.g., happiness, disgust) signifying the maximum “intensity” (a FACS term) of a weighted expression.

EVALUATION OF FERI

In its completed form, the FERI will consist of two basic components. During the first part of the FERI, participants will be shown a series of facial expressions. Within these faces one third will be matched against the participant on race and gender, another third will be matched on race but of the opposite gender, and the last third will be a combination of mismatched races and genders. Additionally, each group will have half high-intensity expressions and half low-intensity expressions. Participants will be asked to identify the presented facial expression, first in an open-choice format and then in a multiple-choice format. Complete verbal responses will be recorded and coded for qualitative information. Children will then be asked to rate their confidence in their response. During the second part of the FERI, participants will be asked to move a slider bar from the starting point of a face’s neutral expression to a requested target emotion. Participants will be asked to stop the slider bar once the target emotion is perceived. Similar to Blair, et al. (2001), the point at which the bars are stopped will be quantified.

Before we integrate the instrument into our studies, we will pilot it to evaluate methodology and ease of use with this population, and its validity against the DANVA2. Participants will include both childhood cancer survivors and healthy controls aged 10 to 16. We will verify that children in this age range can understand the task instructions and successfully manipulate the interface (e.g., a slider bar). We also will obtain a range of time that children require to complete the task. Upon completion of the pilot phase, participants will also be asked if they would be interested in returning within three months to retry the instrument; this reassessment will allow for the calculation of test-retest reliability. Finally, we will assure that there is adequate variability in performance (i.e., that the task is neither too difficult nor too easy for children to complete). Based on these data, the instrument’s user interface will be revised as needed.

All pilot-testing study procedures involving child participants and a parent will occur during one 90-minute session. Participants will complete a brief assessment of their ability to use the mouse to navigate through task-related procedures (e.g., using a slider bar, clicking on multiple-choice responses) and receive additional assistance or training when needed. Then, their general response time will be measured by presenting a series of faces similar to those that will follow in the actual FERI. Instead of requesting children to identify facial expressions, however, children will be asked to identify the gender of each face depicted as quickly as they can without making mistakes. Additionally, a brief visual acuity task will be administered to verify that participants do not have visual impairments that may prevent them from seeing the faces on the task.

To evaluate the adaptability, feasibility, and initial validity of the FERI, we will analyze descriptive and summary statistics of participant-reported and experimenter observations of ease of use, problems, or complaints. Adaptability of the measure will be assessed in terms of an adequate range of performance accuracy across both subject groups (i.e., survivors and controls), expressions (e.g., neutral, sad, happy), and intensity thresholds (e.g., high, low). We also hypothesize that the digital facial expression recognition task will be valid and reliable for use with childhood cancer survivors and healthy children. For these analyses, we will estimate internal consistency using coefficient alpha, and obtain test-retest reliability using data from participants who have agreed to complete the FERI again after a three-month interval. We will assess convergent validity by correlating scores from the forced-choice portions of the instrument with the child faces subtest of the DANVA2. Finally, we hypothesize that errors on the facial expression recognition task will be associated with increased impairment on parent and self-reported measures of social functioning and quality of life. We will look at total errors on the FERI as a predictor of social functioning outcomes in a series of regression analyses.

LIMITATIONS AND FUTURE DIRECTIONS

We see this work as extending to other clinical areas. For instance, we plan to extend our investigation of facial expression of emotion to include broader non-verbal cues, such as body posture or head and hand gestures, all easy to implement with precision with well-understood modifications to the tool. Such an extension would allow researchers to gain a clearer and more complete picture of the deficits faced by childhood cancer survivors and would lead the way towards development of interventions to ameliorate some of these deficits and improved quality of life. Similarly, if expected deficits are found during hypothesis testing (e.g., survivors make significantly more errors in facial recognition than healthy controls; survivors have a higher threshold for perceiving negatively-valenced emotions than healthy controls), then the tool could be used to create a social functioning intervention. One intervention idea would be to create semi-transparent "overlay" templates to train childhood cancer survivors to pay attention to specific, salient features of facial expression. Such a task would allow survivors to become both more efficient and more accurate in their facial expression recognition. Additionally, the tool would allow for the creation of virtual and interactive social scenarios (see Paschall, et al., 2005; Hubal, et al., 2008) that would allow for tutored training and exposure to typical social scenarios that could be practiced until the child is both adept and confident in his/her actions. This tool extension would be relevant for other clinical populations as well, for instance as a training tool for social functioning used by people with autistic spectrum disorders.

ACKNOWLEDGEMENTS

The work described here was performed under award 1-R03-CA-128472-01 from the National Cancer Institute (NCI). Preparation of this paper was partly supported under award IIS-0534211 from the National Science Foundation (NSF). Points of view in this document are those of the authors, and do not necessarily represent the official position of NCI or NSF.

REFERENCES

Blair, R., Colledge, E., Murray, L., & Mitchell, D. (2001). A selective impairment in the processing of sad and fearful expressions in children with psychopathic tendencies. Journal of Abnormal Child Psychology, 29, 491-498.

Blair, R.J.R. (2003). Facial expressions, their communicatory functions and neuro-cognitive substrates. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 358, 561-572.

Bonner, M.J., Hardy, K.K., Willard, V.W, Anthony, K.K., Hood, M., & Gururangan, S. (in press). Social functioning and facial expression recognition in survivors of pediatric brain tumors. Journal of Pediatric Psychology.

Buono, L.A., Morris, M.K., Morris, R.D., Krawiecki, N., Norris, F.H., Foster, M.A., & Copeland, D.R. (1998). Evidence for the syndrome of nonverbal learning disabilities in children with brain tumors. Child Neuropsychology, 4, 144-157.

Carey, M.E., Barakat, L.P., Foley, B., Gyato, K., & Phillips, P.C. (2001). Neuropsychological functioning and social functioning of survivors of pediatric brain tumors: Evidence of a nonverbal learning disability. Child Neuropsychology, 7, 265-272.

Copeland, D.R., Moore, B.D., Francis, D.J., Jaffe, N., & Culbert, S.J. (1996). Neuropsychologic effects of chemotherapy on children with cancer: A longitudinal study. Journal of Clinical Oncology, 14, 2826-2835.

Coupland, N.J., Sustrik, R.A., Ting, P., Li, D., Hartfeil, M., Singh, A.J., & Blair, R.J. (2004). Positive and negative affect differentially influence identification of facial emotions. Depression and Anxiety, 19, 31-34.

Crick, N.R., & Dodge, K.A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. Psychological Bulletin, 115, 74-101.

Ekman, P., Friesen, W.V., & Hager, J.C. (2002). The Facial Action Coding System (2d Edition). London, England: Weidenfeld & Nicolson.

George, P.A., Scaife, M., & Hole, G. J. (2000). Factors influencing young children's ability to discriminate unfamiliar faces by age. International Journal of Behavioural Development, 24, 480-491.

Hubal, R., Kizakevich, P., & Furberg, R. (2007). Synthetic characters in health-related applications. In S. Vaidya, L.C. Jain, & H. Yoshida (Eds.), Advanced Computational Intelligence Paradigms in Healthcare 2 (pp. 5-26). Heidelberg, Germany: Springer-Verlag.

Hubal, R.C., Fishbein, D.H., Sheppard, M.S, Eldreth, D.L., & Hyde, C.T. (2008). How do varied populations interact with embodied conversational agents? Findings from inner-city adolescents and prisoners. Computers in Human Behavior, 24, 1104-1138.

Knapp, M. (1972). Nonverbal Communication in Human Interaction. New York, NY: Reinhart and Winston Inc.

Lemerise, E.A., & Arsenio, W.F. (2000). An integrated model of emotion processes and cognition in social information processing. Child Development, 71, 107-118.

Maddrey, A.M., Bergeron, J.A., Lombardo, E.R., McDonald, N.K., Mulne, A.F., Barenberg, P.D., & Bowers, D.C. (2005). Neuropsychological performance and quality of life of 10 year survivors of childhood medulloblastoma. Journal of Neuro-Oncology, 72, 245-253.

Massaro, D.W., & Bosseler, A. (2006). Read my lips: The importance of the face in a computer-animated tutor for autistic children learning language. Autism: The International Journal of Research and Practice, 10, 495-510.

Mehrabian, A. (1971). Silent Messages. Belmont, CA: Wadsworth.

Moore, B.D. (2005). Neurocognitive outcomes in survivors of childhood cancer. Journal of Pediatric Psychology, 30, 51-63.

Mulhern, R.K., & Palmer, S.L. (2003). Neurocognitive late effects in pediatric cancer. Current Problems in Cancer, 27, 177-197.

Nowicki, S., & Duke, M. (2001). Nonverbal receptivity: The diagnostic analysis of nonverbal accuracy (DANVA). In J.A. Hall & F.J. Bernieri (Eds.), Interpersonal Sensitivity Theory and Measurement (pp. 183-198). Mahwah, NJ: Lawrence Erlbaum Associates.

Paschall, M.J., Fishbein, D.H., Hubal, R.C., & Eldreth, D. (2005). Psychometric properties of virtual reality vignette performance measures: A novel approach for assessing adolescents' social competency skills. Health Education Research: Theory and Practice, 20, 61-70.

Pollack, S.D., & Sinha, P. (2002). Effects of early experience on children's recognition of facial displays of emotion. Developmental Psychology, 38, 784-791.

Schwarzer, G., & Massaro, D.W. (2001). Modeling face identification processing in children and adults. Journal of Experimental Child Psychology, 79, 139-161.

Zebrack, B.J., Gurney, J.G., Oeffinger, K., Whitton, J., Packer, R.J., Mertens, A., Turk, N., Castleberry, R., Dreyer, Z., Robison, L.L., & Zeltzer, L.K. (2004). Psychological outcomes in long-term survivors of childhood brain cancer: A report from the Childhood Cancer Survivors Study. Journal of Clinical Oncology, 22, 999-1006.

Zebrack, B.J., Zeltzer, L.K., Whitton, J., Mertens, A.C., Odom, L., Berkow, R., & Robison, L.L. (2002). Psychological outcomes in long-term survivors of childhood Leukemia, Hodgkin's Disease, and Non-Hodgkin's Lymphoma: A report from the Childhood Cancer Survivor Study. Pediatrics, 110, 42-52.