

Poster

Toward Expert Systems in Mental Health Assessment: A Computational Approach to the Face and Voice in Dyadic Patient-Doctor Interactions

Justin T Baker^{1,2}, MD, PhD; Luciana Pennant¹; Tadas Baltrušaitis³, PhD; Supriya Vijay³; Elizabeth S Liebson^{1,2}, MD; Dost Ongur^{1,2}, MD, PhD; Louis-Philippe Morency³, PhD

¹Psychotic Disorders Division, McLean Hospital, Belmont, MA, United States

²Department of Psychiatry, Harvard Medical School, Boston, MA, United States

³Language Technical Institute, School of Computer Science, Carnegie Mellon University, Pittsburg, PA, United States

Corresponding Author:

Justin T Baker, MD, PhD

Department of Psychiatry

Harvard Medical School

25 Shattuck St

Boston, MA, 02115

United States

Phone: 1 617 855 3913

Fax: 1 617 855 0000

Email: jtbaker@partners.org

Abstract

Background: Computational approaches to measure naturalistic behavior in clinical settings could provide an objective backstop for mental health assessment and disease monitoring, both of which are costly and unreliable using traditional methods.

Objective: The objective of this pilot study was to determine which parts of the mental status exam could be reliably predicted by a combination of facial and vocal features extracted from a recorded interview using a combination of computer-assisted methods, in order to assess feasibility of our approach to quantify behavior for a longitudinal study of patients receiving psychiatric treatment.

Methods: A total of 18 patients carrying diagnoses of schizophrenia, bipolar disorder, and related conditions were recruited from an inpatient psychiatric unit and participated in a total of 24 semi-structured interviews lasting 5-15 minutes (modeled after clinical rounds). Synchronized audio and video data were acquired from both patient and doctor during each encounter using 1080p webcams focused on the face and upper torso and cardioid headset microphones. Standardized psychiatric symptom scales was obtained after each recorded interview. Behavioral features, including facial action units (AUs), gaze, and speech characteristics (eg, prosody, pitch, tone, texture) were computed automatically using in-house and publicly available software. To predict clinical scales we trained a linear kernel support vector regressor (SVR) using features from both the entire session (ie, global mean) and each experimental epoch (eg, means during time spent alone and each individual question), leading to 15 predictors for each clinical scale item and scale totals. We used leave-one-out validation on the training data (maximizing the Pearson correlation coefficient) to determine the C parameter for the SVR models; for testing, we used leave-one-subject-out cross-validation (ie, leaving 17 participants for training/validation in each fold).

Results: Providing evidence of our approach's ability to capture and quantify relevant signal that confirms or verifies clearly visible psychopathology, we found that parameters such as brow furrowing (AU4, $R=0.744$) and eye widening (AU5, $R=-0.601$) were correlated with depression measures on the BPRS. In many cases, these effects were specific to the question or experimental epoch. For instance, unusual thought content was most evident in increased frequency of brow flashes (AU2, $R=0.752$) and greater smile variability ($R=0.656$) that occurred while participants were alone in the room. Individuals with higher ratings of delusions also showed increased brow flashes in response to a question about their self confidence ($R=0.739$). Many relationships showed a "dose effect" with midrange scores corresponding with moderate psychopathology.

Conclusions: Our experiments show that automatically detected facial action units and speech properties can be used to predict and quantify a number of psychiatric symptoms from multiple domains of psychopathology, including both mood and psychosis.

We demonstrate the importance of analyzing behaviors in the appropriate context (ie, while participants are alone or prompted with a specific question) in order to optimally extract clinically relevant information from objective indices of behavior. Thus, quantitative assessment of behavior in naturalistic settings is both feasible and informative as an adjunct to traditional methods of mental status assessment.

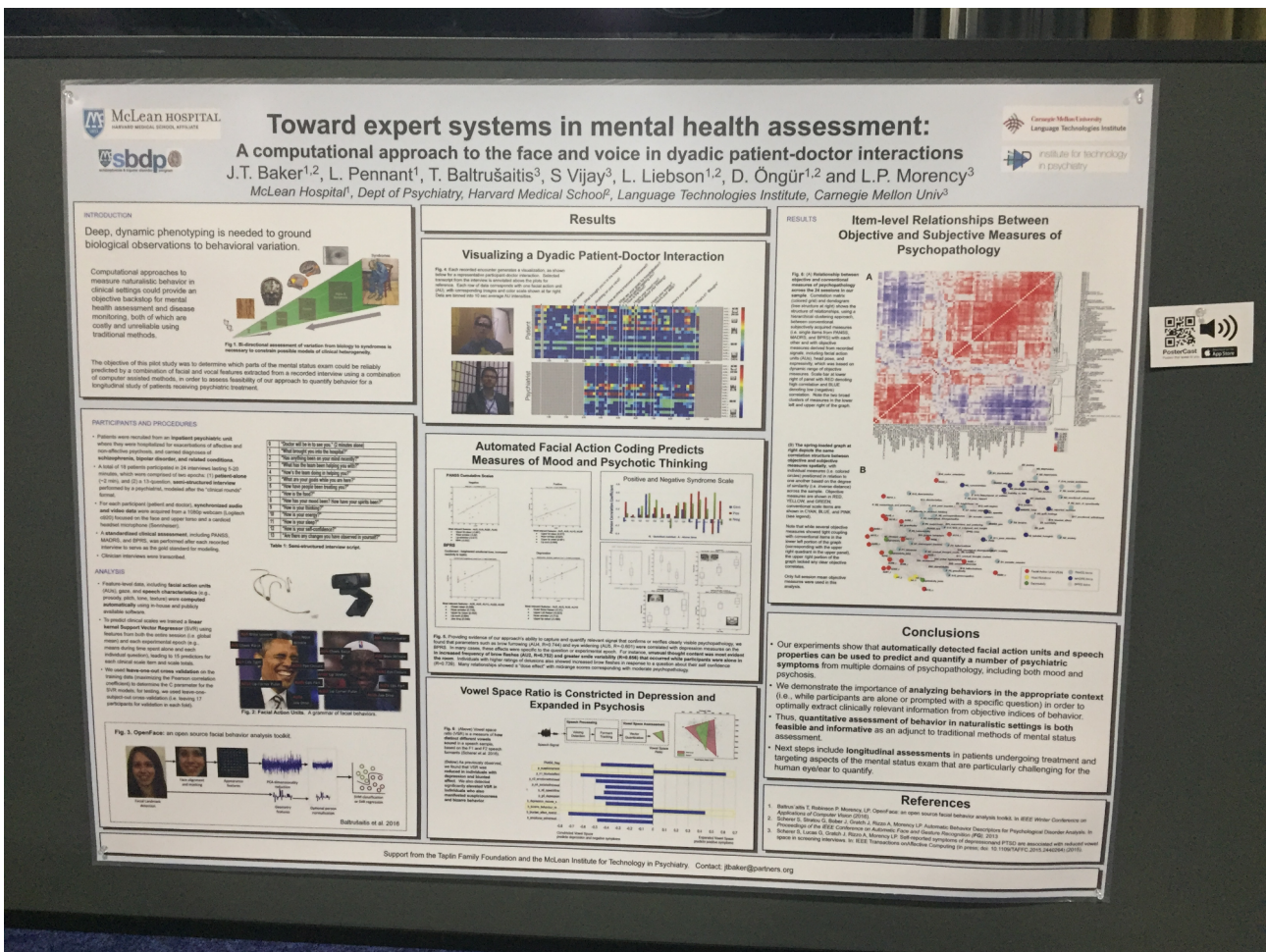
(iproc 2016;2(1):e44) doi: 10.2196/iproc.6136

KEYWORDS

human interaction; facial expression; voice; depression; bipolar disorder; schizophrenia; patient-physician relationship

This poster was presented at the Connected Health Symposium 2016, October 20-21, Boston, MA, United States. A photo of the poster is displayed as an image in Figure 1 and as a higher resolution image in Multimedia Appendix 1.

Figure 1. Poster.



Multimedia Appendix 1

Poster.

[JPG File, 2MB-Multimedia Appendix 1]

Edited by T Hale; submitted 05.06.16; peer-reviewed by CHS Scientific Program Committee; accepted 02.08.16; published 30.12.16

Please cite as:

Baker JT, Pennant L, Baltrušaitis T, Vijay S, Liebson ES, Ongur D, Morency LP

Toward Expert Systems in Mental Health Assessment: A Computational Approach to the Face and Voice in Dyadic Patient-Doctor Interactions

iproc 2016;2(1):e44

URL: <http://www.iproc.org/2016/1/e44/>

doi: [10.2196/iproc.6136](https://doi.org/10.2196/iproc.6136)

PMID:

©Justin T Baker, Luciana Pennant, Tadas Baltrušaitis, Supriya Vijay, Elizabeth S Liebson, Dost Ongur, Louis-Philippe Morency. Originally published in Iproceedings (<http://www.iproc.org>), 30.12.2016. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in Iproceedings, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.iproc.org/>, as well as this copyright and license information must be included.