

Nonverbal Social Withdrawal in Depression: Evidence from manual and automatic analysis

Jeffrey M. Girard^{a,*}, Jeffrey F. Cohn^{a,b}, Mohammad H. Mahoor^c, S. Mohammad Mavadati^c,
Zakia Hammal^b, Dean P. Rosenwald^a

^a4322 Sennott Square, Department of Psychology, University of Pittsburgh, Pittsburgh, PA, USA 15260

^b5000 Forbes Avenue, The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, USA, 15213

^c2390 S. York Street, Department of Electrical and Computer Engineering, University of Denver, Denver, CO, USA, 80208

Abstract

The relationship between nonverbal behavior and severity of depression was investigated by following depressed participants over the course of treatment and video recording a series of clinical interviews. Facial expressions and head pose were analyzed from video using manual and automatic systems. Both systems were highly consistent for FACS action units (AUs) and showed similar effects for change over time in depression severity. When symptom severity was high, participants made fewer affiliative facial expressions (AUs 12 and 15) and more non-affiliative facial expressions (AU 14). Participants also exhibited diminished head motion (i.e., amplitude and velocity) when symptom severity was high. These results are consistent with the Social Withdrawal hypothesis: that depressed individuals use nonverbal behavior to maintain or increase interpersonal distance. As individuals recover, they send more signals indicating a willingness to affiliate. The finding that automatic facial expression analysis was both consistent with manual coding and revealed the same pattern of findings suggests that automatic facial expression analysis may be ready to relieve the burden of manual coding in behavioral and clinical science.

Keywords: Depression, Multimodal, FACS, Facial Expression, Head Motion

1. Introduction

Unipolar depression is a common psychological disorder and one of the leading causes of disease burden worldwide [56]; its symptoms are broad and pervasive, impacting affect, cognition, and behavior [2]. As a channel of both emotional expression and interpersonal communication, nonverbal behavior is central to how depression presents and is maintained [1, 51, 76]. Several theories of depression address nonverbal behavior directly and make predictions about what patterns should be indicative of the depressed state.

The Affective Dysregulation hypothesis [15] interprets depression in terms of *valence*, which captures the positivity or negativity (i.e., pleasantness or aversiveness) of an emotional state. This hypothesis argues that depression is marked by deficient positivity and excessive negativity. Deficient positivity is consistent with many neurobiological theories of depression [22, 42] and highlights the symptom of anhedonia. Excessive negativity is consistent with many cognitive theories of depression [5, 8] and highlights the symptom of rumination. In support of the Affective Dysregulation hypothesis, with few exceptions [61], observational studies have found that depression is marked by reduced positive expressions, such as smiling and laughter [7, 13, 26, 38, 39, 47, 49, 62, 64, 66, 67, 75, 78, 80, 83]. Several studies also found that depression is marked by increased negative expressions [9, 26, 61, 74]. However, other studies found the opposite effect: that

*Corresponding author +1 412 624 8826

Email addresses: jmg174@pitt.edu (Jeffrey M. Girard), jeffc@cs.cmu.edu (Jeffrey F. Cohn), mohammad.mahoor@du.edu (Mohammad H. Mahoor), seyedmohammad.mavadati@du.edu (S. Mohammad Mavadati), zakia_hammal@yahoo.fr (Zakia Hammal), dpr10@pitt.edu (Dean P. Rosenwald)

depression is marked by reduced negative expressions [38, 47, 62]. These studies, combined with findings of reduced emotional reactivity using self-report and physiological measures (see Bylsma et al. [10] for a review), led to the development of an alternative hypothesis.

The Emotion Context Insensitivity hypothesis [63] interprets depression in terms of deficient *appetitive motivation*, which directs behavior towards exploration and the pursuit of potential rewards. This hypothesis argues that depression is marked by a reduction of both positive and negative emotion reactivity. Reduced reactivity is consistent with many evolutionary theories of depression [34, 52, 60] and highlights the symptoms of apathy and psychomotor retardation. In support of this hypothesis, some studies found that depression is marked by reductions in general facial expressiveness [26, 38, 47, 62, 69, 79] and head movement [30, 48]. However, it is unclear how much of this general reduction is accounted for by reduced positive expressions, and it is problematic for the hypothesis that negative expressions are sometimes increased in depression.

As both hypotheses predict reductions in positive expressions, their main distinguishing feature is their treatment of negative facial expressions. The lack of a clear result on negative expressions is thus vexing. Three limitations of previous studies likely contributed to these mixed results.

First, most studies compared the behavior of depressed participants with that of non-depressed controls. This type of comparison is problematic because depression is highly correlated with numerous personality traits that influence nonverbal behavior. For instance, people with psychological disorders in general are more likely to have high neuroticism (more anxious, moody, and jealous) and people with depression in particular are more likely to have low extraversion (less outgoing, talkative, and energetic) [54]. Thus, many previous studies confounded depression with the stable personality traits that are correlated with it.

Second, many studies used experimental contexts of low sociality, such as viewing emotional stimuli while alone. However, many nonverbal behaviors - especially those involved in social signaling - are typically rare in such contexts [36]. Thus, many previous studies may have introduced a “floor effect” by using contexts in which relevant expressions were unlikely to occur in any of the experimental groups.

Finally, most studies examined a limited range of nonverbal behavior. Many aggregated multiple facial expressions into single categories such as “positive expressions” and “negative expressions,” while others used single facial expressions to represent each category (e.g., smiles for positive affect and brow-lowering for negative affect). These approaches assume that all facial expressions in each category are equivalent and will be affected by depression in the same way. However, in addition to expressing a person’s affective state, nonverbal behavior is also capable of regulating social interactions and communicating behavioral intentions [36, 37]. Perhaps individual nonverbal behaviors are increased or decreased in depression based on their social-communicative function. Thus, we propose an alternative hypothesis to explain the nonverbal behavior of depressed individuals, which we refer to as the Social Withdrawal hypothesis.

The Social Withdrawal hypothesis interprets depression in terms of *affiliation*: the motivation to cooperate, comfort, and request aid. This hypothesis argues that depression is marked by reduced affiliative behavior and increased non-affiliative behavior, which combine to maintain or increase interpersonal distance. This pattern is consistent with several evolutionary theories [1, 76] and highlights the symptom of social withdrawal. It also accounts for variability within the group of negative facial expressions, as each varies in its social-communicative function. Specifically, expressions that signal approachability should be decreased in depression and expressions that signal hostility should be increased. In support of this hypothesis, several ethological studies of depressed inpatients found that their nonverbal behavior was marked by hostility and a lack of social engagement [23, 33, 67, 79]. Findings of decreased smiling in depression also support this interpretation, as smiling is often an affiliative signal [43, 44, 53].

1.1. The Current Study

We address many of the limitations of previous work by: (1) following depressed participants over the course of treatment; (2) observing them during a clinical interview; and (3) measuring multiple, objectively-defined nonverbal behaviors. By comparing participants to themselves over time, we control for stable personality traits; by using a clinical interview, we provide a context that is both social and representative of typical client-patient interactions; and by examining head motion and multiple facial movements individually, we explore the possibility that depression may affect different aspects of nonverbal behavior in different ways.

We measure nonverbal behavior with both manual and automatic coding. In an effort to alleviate the substantial time burden of manual coding, interdisciplinary researchers have begun to develop automatic systems to supplement (and perhaps one day replace) human coders [84]. Initial applications to the study of depression are also beginning to appear [17, 48, 49, 59, 68, 69]. However, many of these applications are “black-box” procedures that are difficult to interpret, some require the use of invasive recording equipment, and most have not been validated against expert human coders.

We examine patterns of head motion and the occurrence of four facial expressions that have been implicated as affiliative or non-affiliative signals (Figure 1). These expressions are defined by the Facial Action Coding System (FACS) [24] in terms of individual muscle movements called action units (AUs).

First, we examine the lip corner puller or smile expression (AU 12). As part of the prototypical expression of happiness, the smile is implicated in positive affect [25] and may signal affiliative intent [43, 44, 53]. Second, we examine the dimpler expression (AU 14). As part of the prototypical expression of contempt, the dimpler is implicated in negative affect [25] and may signal non-affiliative intent [43]. The dimpler can also counteract and obscure an underlying smile, potentially augmenting its affiliative signal and expressing embarrassment or ambivalence [26, 50]. Third, we examine the lip corner depressor expression (AU 15). As part of the prototypical expression of sadness, the lip corner depressor is implicated in negative affect [25] and may signal affiliative intent (especially the desire to show or receive empathy) [44]. Fourth, we examine the lip presser expression (AU 24). As part of the prototypical expression of anger, the lip presser is implicated in negative affect [25] and may signal non-affiliative intent [43, 44, 53]. Finally, we also examine the amplitude and velocity of head motion. Freedman [35] has argued that movement behavior should be seen as part of an effort to communicate. As such, head motion may signal affiliative intent.

We hypothesize that the depressed state will be marked by reduced amplitude and velocity of head motion, reduced activity of affiliative expressions (AUs 12 and 15), and increased activity of non-affiliative expressions (AUs 14 and 24). We also test the predictions of two established hypotheses. The Affective Dysregulation hypothesis is that depression will be marked by increased negativity (AUs 14, 15, and 24) and decreased positivity (AU 12), and the Emotion Context Insensitivity hypothesis is that the depressed state will be marked by reduced activity of all behavior (AUs 12, 14, 15, 24, and head motion).

2. Methods

2.1. Participants

The current study analyzed video of 33 adults from a clinical trial for treatment of depression [17]. At the time of study intake, all participants met DSM-IV [2] criteria for major depressive disorder [29]. In the clinical trial, participants were randomly assigned to receive either antidepressant medication (i.e., SSRI) or interpersonal psychotherapy (IPT); both treatments are empirically validated for use in depression [45].

Symptom severity was evaluated on up to four occasions at 1, 7, 13, and 21 weeks by clinical interviewers ($n=9$, all female). Interviewers were not assigned to specific participants and varied in the number of interviews they conducted. Interviews were conducted using the Hamilton Rating Scale for Depression (HRSD) [40], which is a criterion measure for assessing severity of depression. Interviewers all were expert in the HRSD and reliability was maintained above 0.90. HRSD scores of 15 or higher generally indicate moderate to severe depression, and scores of 7 or lower generally indicate a return to normal [65].

Interviews were recorded using four hardware-synchronized analogue cameras. The video data used in the current study came from a camera positioned roughly 15 degrees to the participant’s right; its images were digitized into 640x480 pixel arrays at a frame rate of 29.97 frames per second. Only the first three interview questions (about depressed mood, feelings of guilt, and suicidal ideation) were analyzed. These segments ranged in length from 28 to 242 seconds with an average of 100 seconds.

The automatic facial expression coding system was trained using data from all 33 participants. However, not all participants were included in the longitudinal depression analyses. In order to compare participants’ nonverbal behavior during high and low symptom severity, we only analyzed participants whose symptoms improved over the course of treatment. All participants had high symptom severity ($\text{HRSD} \geq 15$) during their initial interview. “Responders” were those participants who had low symptom severity ($\text{HRSD} \leq 7$)

during a subsequent interview. Although 21 participants met this criterion, data from 3 participants were unable to be processed due to camera occlusion and gum-chewing. The final group was thus composed of 38 interviews from 19 participants; this group’s demographics were similar to those of the database (Table 1).

2.2. Manual FACS Coding

The Facial Action Coding System (FACS) [24] is the current gold standard for facial expression annotation. FACS decomposes facial expressions into component parts called action units (AUs). Action units are anatomically-based facial movements that correspond to the contraction of specific muscles. For example, AU 12 indicates contractions of the zygomatic major muscle and AU 14 indicates contractions of the buccinator muscle. AUs may occur alone or in combination to form more complex expressions.

The facial behavior of all participants was FACS coded from video by certified and experienced coders. Expression onset, apex, and offset were coded for 17 commonly occurring AUs. Overall inter-observer agreement for AU occurrence, quantified by Cohen’s Kappa [16], was 0.75; according to convention, this can be considered good agreement [31]. In order to minimize the number of analyses required, the current study analyzed four of these AUs that are conceptually related to emotion and social signaling (Figure 1).

2.3. Automatic FACS Coding

2.3.1. Face Registration

In order to register two facial images (i.e., the “reference image” and the “sensed image”), a set of facial landmark points is utilized. Facial landmark points indicate the location of important facial components (e.g., eye corners, nose tip, jaw line). Active Appearance Models (AAM) were used to track sixty-six facial landmark points (Figure 2) in the video. AAM is a powerful approach that combines shape and texture variation of an image into a single statistical model that can simultaneously match the face model into a new facial image [19]. Approximately 3% of video frames were manually annotated for each subject to train the AAMs. The tracked images were automatically aligned using a gradient-descent AAM fitting algorithm [57]. In order to enable comparison between images, differences in scale, translation, and rotation were removed using a two dimensional similarity transformation.

2.3.2. Feature Extraction

Facial expressions appear as changes in facial shape (e.g., curvedness of the mouth) and appearance (e.g., wrinkles and furrows). In order to describe facial expression efficiently, it would be helpful to utilize a representation that can simultaneously describe the shape and appearance of an image. Localized Binary Pattern Histogram (LBPH), Histogram of Oriented Gradient (HOG), and Localized Gabor are some well-known features that can be used for this purpose [77]. We used Localized Gabor features for texture representation and discrimination as they have been found to be relatively robust to alignment error [14]. A bank of 40 Gabor filters (i.e., 5 scales and 8 orientations) was applied to regions defined around the 66 landmark points, which resulted in 2640 Gabor features per frame.

2.3.3. Dimensionality Reduction

In many pattern classification applications, the number of features is extremely large. Their high dimensionality makes analysis and classification a complex task. In order to extract the most important or discriminant features of the data, several linear techniques, such as Principle Component Analysis (PCA) and Independent Component Analysis (ICA), and nonlinear techniques, such as Kernel PCA and Manifold Learning, have been proposed [32]. We utilized the manifold learning technique to reduce the dimensionality of the Gabor features. The idea behind this technique is that the data points lie on a low dimensional manifold that is embedded in a high dimensional space [11]. Specifically, we utilized Laplacian Eigenmap [6] to extract the low-dimensional features and then, similar to Mahoor et al. [55], we applied spectral regression to calculate the corresponding projection function for each AU. Through this step, the dimensionality of the Gabor data was reduced to 29 features per frame.

2.3.4. Classifier Training

In machine learning, feature classification aims to assign each input value to one of a given set of classes. A well-known classification technique is the Support Vector Machine (SVM). The SVM classifier applies the “kernel trick,” which uses dot product, to keep computational loads reasonable. The kernel functions (e.g., linear, polynomial, and radial basis function) enable the SVM algorithm to fit a hyperplane with a maximum margin into the transformed high dimensional feature space. Radial basis function kernels were used in the current study. To build a training set, we randomly sampled positive and negative frames from each subject. To train and test the SVM classifiers, we used the LIBSVM library [12]. To find the best classifier and kernel parameters (i.e., C and γ , respectively), we used a “grid-search” procedure during leave-one-out cross-validation [46].

2.4. Automatic Head Pose Analysis

A cylinder-based 3D head tracker (CSIRO tracker) [20] was used to model rigid head motion (Figure 3). Angles of the head in the horizontal and vertical directions were selected to measure head pose and motion. These directions correspond to the meaningful motion of head nods and head turns. In a recent behavioral study, the CSIRO tracker demonstrated high concurrent validity with alternative methods for head tracking and revealed differences in head motion between distressed and non-distressed intimate partners [41].

For each video frame, the tracker output six degrees of freedom of rigid head motion or an error message that the frame could not be tracked. Overall, 4.61% of frames could not be tracked. To evaluate the quality of the tracking, we visually reviewed the tracking results overlaid on the video. Visual review indicated poor tracking in 11.37% of frames; these frames were excluded from further analysis. To prevent biased measurements due to missing data, head motion measures were computed for the longest segment of continuously valid frames in each interview. The mean duration of these segments was 55.00 and 37.31 seconds for high and low severity interviews, respectively.

For the measurement of head motion, head angles were first converted into amplitude and velocity. Amplitude corresponds to the amount of displacement, and velocity (i.e., the derivative of displacement) corresponds to the speed of displacement for each head angle during the segment. Amplitude and velocity were converted into *angular* amplitude and angular velocity by subtracting the mean overall head angle across the segment from the head angle for each frame. Because the duration of each segment was variable, the obtained scores were normalized by the duration of its corresponding segment.

2.5. Measures

To capture how often different facial expressions occurred, we calculated the base rate for each AU. Base rate was calculated as the number of frames for which the AU was present divided by the total number of frames. This measure was calculated for each AU as determined by both manual and automatic FACS coding. To represent head motion, average angular amplitude and velocity were calculated using root mean square, which is a measure of the magnitude of a varying quantity.

2.6. Data Analysis

The primary analyses compared the nonverbal behavior of participants at two time points: when they were severely depressed and when they had recovered. Because base rates for the AU were highly skewed (each $p < 0.05$ by Shapiro-Wilk test [72]), differences with severity are reported for non-parametric tests: the Wilcoxon Signed Ranks test [82]. Results using parametric tests (i.e., paired t-tests) were comparable.

We compared manual and automatic FACS coding in two ways. First, we measured the degree to which both coding sources detected the same AUs in each frame. This frame-level reliability was quantified using area under the ROC curve (AUC) [28], which corresponds to the automatic system’s discrimination (the likelihood that it will correctly discriminate between a randomly-selected positive example and a randomly-selected negative example). Additionally, we measured the degree to which the two methods were consistent in measuring base rate for each session. This session-level reliability was quantified using intraclass correlation (ICC) [73], which describes the resemblance of units in a group.

3. Results

Manual and automatic FACS coding demonstrated high frame-level and session-level reliability (Table 2) and revealed similar patterns between high and low severity interviews (Table 3). Manual analysis found that two AUs were reduced during high severity interviews (AUs 12 and 15) and one facial expression was increased (AU 14); automatic analysis replicated these results except that the difference with AU 12 did not reach statistical significance. Automatic head pose analysis also revealed significant differences between interviews. Head amplitude and velocity were decreased in high severity interviews (Table 4).

4. Discussion

Facial expression and head motion tracked changes in symptom severity. They systematically differed between times when participants were severely depressed (Hamilton score of 15 or greater) and those during which depression abated (Hamilton score of 7 or less). While some specific findings were consistent with all three alternative hypotheses, support was strongest for Social Withdrawal.

According to the Affective Dysregulation hypothesis, depression entails a deficit in the ability to experience positive emotion and enhanced potential for negative emotion. Consistent with this hypothesis, when symptom severity was high, AU 12 occurred less often and AU 14 occurred more often. Participants smiled less and showed expressions related to contempt and embarrassment more often when severity was high. When depression abated, smiling increased and expressions related to contempt and embarrassment decreased.

There were, however, two findings contrary to Affective Dysregulation. One, AU 24, which is associated with anger, failed to vary with depression severity. Because Affective Dysregulation makes no predictions about specific emotions, this negative finding might be discounted. Anger may not be among the negative emotions experienced in interviews. A second contrary finding is not so easily dismissed. AU 15, which is associated with both sadness and positive affect suppression, was decreased when depression was severe and increased when depression abated. An increase in negative emotion when depression abates runs directly counter to Affective Dysregulation. Thus, while the findings for AU 12 and AU 14 were supportive of Affective Dysregulation, the findings for AU 15 were not. Moreover, the findings for head motion are not readily interpreted in terms of this hypothesis. Because head motion could be related to either positive or negative emotion, this finding is left unaccounted for. Thus, while some findings were consistent with the hypothesis, others importantly were not.

According to the Emotion Context Insensitivity hypothesis, depression entails a curtailment of both positive and negative emotion reactivity. Consistent with this hypothesis, when symptom severity was high, AU 12 and AU 15 occurred less often. Participants smiled less and showed expressions related to sadness less often when severity was high. The decreased head motion in depression was consistent with the lack of reactivity assumed by the hypothesis. When depression abated, smiling, expressions related to sadness, and head motion all increased. These findings are consistent with Emotion Context Insensitivity.

However, two findings ran contrary to this hypothesis. AU 24 failed to vary with depression level, and AU 14 actually increased when participants were depressed. While the lack of findings for AU 24 might be dismissed, for reasons noted above, the results for AU 14 are difficult to explain from the perspective of Emotion Context Insensitivity. In previous research, support for this hypothesis has come primarily from self-report and physiological measures [10]. Emotion Context Insensitivity could reveal a lack of coherence among them. Dissociations between expressive behavior and physiology have been observed in other contexts [58]; depression may be another. Further research will be needed to explore this question.

The Social Withdrawal hypothesis emphasizes behavioral intention to affiliate rather than affective valence. According to this hypothesis, affiliative behavior decreases and non-affiliative behavior increases when severity is high. Decreased head motion together with decreased occurrence of AUs 12 and 15 and increased occurrence of AU 14 all were consistent with this hypothesis. When depression severity was high, participants smiled less, showed expressions related to sadness less often, showed expressions related to contempt and embarrassment more often, and showed smaller and slower head movements. When depression abated, smiling, expressions related to sadness, and head motion increased, while expressions related to contempt

and embarrassment decreased. The only mark against this hypothesis is the failure to find a change in AU 24. As a smile control, AU 24 may occur less often than AU 14 and other smile controls. In a variety of contexts, AU 14 has been the most noticed smile control [27, 61]. Little is known, however, about the relative occurrence of different smile controls.

These results suggest that social signaling is critical to understanding the link between nonverbal behavior and depression. When depressed, participants signaled intention to avoid or minimize affiliation. Displays of this sort may contribute to the resilience of depression. It has long been known that depressed behavior is experienced as aversive by non-depressed persons [21]. Expressive behavior in depression may discourage social interaction and support that are important to recovery. Ours may be the first study to reveal specific displays responsible for the relative isolation that depression achieves.

Previous work called attention to social skills deficits [70] as the source of interpersonal rejection in depression [71]. According to this view, depressed persons are less adept at communicating their wishes, needs, and intentions with others. This view is not mutually exclusive with the Social Withdrawal hypothesis. Depressed persons might be less effective in communicating their wants and desires and less effective in resolving conflict, yet also signal lack of intention to socially engage with others to accomplish such goals. From a treatment perspective, it would be important to assess which of these factors were contributing to the social isolation experienced in depression.

The automatic facial expression analysis system was consistent with manual coding in terms of frame-level discrimination. The average AUC score across all AUs was 0.86, which is comparable with or better than other studies that detected AUs in spontaneous video [4, 18, 81, 85]. The automatic and manual systems were also highly consistent in terms of measuring the base rate of each AU in the interview session; the average ICC score across all AUs was 0.91. Furthermore, the significance tests using automatic coding were very similar to those using manual coding. The only difference was that the test of AU 12 base rate was significant for manual but not automatic coding, although the findings were in the same direction. To our knowledge, this is the first study to demonstrate that automatic and manual FACS coding are fungible; automatic and manual coding reveal the same patterns. These findings suggest automatic coding may be a valid alternative to manual coding. If so, this would free research from the severe constraints of manual coding. Manual coding requires hours for a person to label a minute or so of video, and that is after what is already a lengthy period of about 100 hours to gain competence in FACS.

The current study represents a step towards comprehensive description of nonverbal behavior in depression (see also [26, 38]). Analyzing individual AUs allowed us to identify a new pattern of results with theoretical and clinical implications. However, this fine-grained analysis also makes interpretation challenging because we cannot definitively say what an individual AU means. We address this issue by emphasizing the relation of AUs to broad dimensions of behavior (i.e., valence and affiliation). We also selected AUs that are relatively unambiguous. For example, we did not include the brow lowerer (AU 4) because it is implicated in fear, sadness, and anger [25] as well as concentration [3] and is thus ambiguous.

In conclusion, we found that high depression severity was marked by a decrease in specific nonverbal behaviors (AU 12, AU 15, and head motion) and an increase in at least one other (AU 14). These results help to parse the ambiguity of previous findings regarding negative facial expressions in depression. Rather than a global increase or decrease in such expressions, these results suggest that depression affects them differentially based on their social-communicative value: affiliative expressions are reduced in depression and non-affiliative expressions are increased. This pattern supports the Social Withdrawal hypothesis: that nonverbal behavior in depression serves to maintain or increase interpersonal distance, thus facilitating social withdrawal. Additionally, we found that the automatic coding was highly reliable with manual coding, yielded convergent findings, and appears ready for independent use in clinical research.

Acknowledgments

The authors wish to thank Nicole Siverling and Shawn Zuratovic for their generous assistance. This work was supported in part by US National Institutes of Health grants MH51435, MH65376, and MH096951 to the University of Pittsburgh. Any opinions, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Institutes of Health.

References

- [1] N.B. Allen, P.B.T. Badcock, The Social Risk Hypothesis of Depressed Mood: Evolutionary, Psychosocial, and Neurobiological Perspectives, *Psychological Bulletin* 129 (2003) 887–913.
- [2] American Psychiatric Association, *Diagnostic and statistical manual of mental disorders* (4th ed.), Washington, DC, 1994.
- [3] S. Baron-Cohen, *Mind reading: The interactive guide to emotion*, Jessica Kingsley Publishers, London, 2003.
- [4] M.S. Bartlett, G. Littlewort, M.G. Frank, C. Lainscsek, I.R. Fasel, J.R. Movellan, Automatic Recognition of Facial Actions in Spontaneous Expressions, *Journal of Multimedia* 1 (2006) 22–35.
- [5] A.T. Beck, *Depression: Clinical, experimental, and theoretical aspects*, Harper and Row, New York, 1967.
- [6] M. Belkin, P. Niyogi, Laplacian eigenmaps for dimensionality reduction and data representation, *Neural Computation* 15 (2003) 1373–1396.
- [7] H. Berenbaum, T.F. Oltmanns, Emotional experience and expression in schizophrenia and depression, *Journal of Abnormal Psychology* 101 (1992) 37–44.
- [8] G.H. Bower, Mood and memory, *American Psychologist* 36 (1981) 129.
- [9] A.Z. Brozgold, J.C. Borod, C.C. Martin, L.H. Pick, M. Alpert, J. Welkowitz, Social functioning and facial emotional expression in neurological and psychiatric disorders, *Applied Neuropsychology* 5 (1998) 15–23.
- [10] L.M. Bylsma, B.H. Morris, J. Rottenberg, A meta-analysis of emotional reactivity in major depressive disorder, *Clinical Psychology Review* 28 (2008) 676–691.
- [11] L. Cayton, Algorithms for manifold learning, Technical Report CS2008-0923, 2005.
- [12] C.C. Chang, C.J. Lin, LIBSVM: a library for support vector machines, *ACM Transactions on Intelligent Systems and Technology* 2 (2011) 27.
- [13] Y.E. Chentsova-Dutton, J.L. Tsai, I.H. Gotlib, Further evidence for the cultural norm hypothesis: positive emotion in depressed and control European American and Asian American women, *Cultural Diversity and Ethnic Minority Psychology* 16 (2010) 284–295.
- [14] S.W. Chew, P. Lucey, S. Lucey, J. Saragih, J.F. Cohn, I. Matthews, S. Sridharan, In the Pursuit of Effective Affective Computing: The Relationship Between Features and Registration, *IEEE Transactions on Systems, Man, and Cybernetics* 42 (2012) 1006–1016.
- [15] L.A. Clark, D. Watson, Tripartite model of anxiety and depression: psychometric evidence and taxonomic implications, *Journal of Abnormal Psychology* 100 (1991) 316.
- [16] J. Cohen, A coefficient of agreement for nominal scales, *Educational and Psychological Measurement* 20 (1960) 37–46.
- [17] J.F. Cohn, T.S. Kruez, I. Matthews, Y. Ying, N. Minh Hoai, M.T. Padilla, Z. Feng, F. De La Torre, Detecting depression from facial actions and vocal prosody, in: *Affective Computing and Intelligent Interaction*, Amsterdam, pp. 1–7.
- [18] J.F. Cohn, M.A. Sayette, Spontaneous facial expression in a small group can be automatically measured: An initial demonstration, *Behavior Research Methods* 42 (2010) 1079–1086. URL: <http://dx.doi.org/10.3758/BRM.42.4.1079>.
- [19] T.F. Cootes, G.J. Edwards, C.J. Taylor, Active appearance models, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23 (2001) 681–685.
- [20] M. Cox, J. Nuevo-Chiquero, J.M. Saragih, S. Lucey, CSIRO Face Analysis SDK, 2013.
- [21] J.C. Coyne, Toward an interactional description of depression, *Psychiatry: Journal for the study of interpersonal processes* (1976).
- [22] R. Depue, W. Iacono, Neurobehavioral aspects of affective disorders, *Annual review of psychology* 40 (1989) 457–492.
- [23] A.K. Dixon, H.U. Fisch, C. Huber, A. Walser, *Ethological Studies in Animals and Man, Their Use in Psychiatry*, *Pharmacopsychiatry* 22 (1989) 44–50.
- [24] P. Ekman, W.V. Friesen, J. Hager, *Facial Action Coding System (FACS): A technique for the measurement of facial movement*, Research Nexus, Salt Lake City, UT, 2002.
- [25] P. Ekman, D. Matsumoto, *Facial expression analysis*, *Scholarpedia* 3 (2008).
- [26] H. Ellgring, *Nonverbal communication in depression*, Cambridge University Press, Cambridge, 1989.
- [27] C.E. Fairbairn, M.a. Sayette, J.M. Levine, J.F. Cohn, K.G. Creswell, The Effects of Alcohol on the Emotional Displays of Whites in Interracial Groups, *Emotion* (Washington, D.C.) 12 (2013).
- [28] T. Fawcett, An introduction to ROC analysis, *Pattern Recognition Letters* 27 (2006) 861–874.
- [29] M.B. First, M. Gibbon, R.L. Spitzer, J.B. Williams, *User’s guide for the structured clinical interview for DSM-IV Axis I disorders (SCID-I, Version 2.0, October 1995 Final Version)*, Biometric Research, New York, 1995.
- [30] H.U. Fisch, S. Frey, H.P. Hirsbrunner, Analyzing nonverbal behavior in depression, *Journal of abnormal psychology* 92 (1983) 307–18.
- [31] J.L. Fleiss, *Statistical methods for rates and proportions*, 2nd ed., John Wiley & Sons, New York, 1981.
- [32] I.K. Fodor, A survey of dimension reduction techniques, Technical Report, 2002.
- [33] L. Fossi, C. Faravelli, M. Paoli, The ethological approach to the assessment of depressive disorders, *Journal of Nervous and Mental Disease* (1984).
- [34] D.C. Fowles, A Motivational View of Psychopathology, in: W.D. Spaulding, H.A. Simon (Eds.), *Integrative Views of Motivation, Cognition, and Emotion*, U of Nebraska Press, 1994, pp. 181–238.
- [35] N. Freedman, The analysis of movement behavior during the clinical interview, in: A.W. Siegman, B. Pope (Eds.), *Studies in dyadic communication*, Pergamon Press, New York, 1972.
- [36] A.J. Fridlund, *Human facial expression: An evolutionary view*, Academic Press, 1994.
- [37] N.H. Frijda, Expression, emotion, neither, or both?, *Cognition and Emotion* 9 (1995) 617–635.
- [38] W. Gabel, W. Wolwer, Facial expressivity in the course of schizophrenia and depression, *European Archives of Psychiatry and Clinical Neuroscience* 254 (2004) 335–342.

- [39] J.G. Gehricke, D. Shapiro, Reduced facial expression and social context in major depression: discrepancies between facial muscle activity and self-reported emotion, *Psychiatry Research* 95 (2000) 157–167.
- [40] M. Hamilton, Development of a rating scale for primary depressive illness, *British Journal of Social and Clinical Psychology* 6 (1967) 278–296.
- [41] Z. Hammal, T.E. Bailie, J.F. Cohn, D.T. George, S. Lucey, Temporal Coordination of Head Motion in Couples with History of Interpersonal Violence, in: *International Conference on Automatic Face and Gesture Recognition*, Shanghai, China.
- [42] J.B. Henriques, R.J. Davidson, Decreased responsiveness to reward in depression, *Cognition & Emotion* 14 (2000) 711–724.
- [43] U. Hess, R.B. Adams Jr, R.E. Kleck, Who may frown and who should smile? Dominance, affiliation, and the display of happiness and anger, *Cognition and Emotion* 19 (2005) 515–536.
- [44] U. Hess, S. Blairy, R.E. Kleck, The influence of facial emotion displays, gender, and ethnicity on judgments of dominance and affiliation, *Journal of Nonverbal Behavior* 24 (2000) 265–283.
- [45] S.D. Hollon, M.E. Thase, J.C. Markowitz, Treatment and Prevention of Depression, *Psychological Science in the Public Interest* 3 (2002) 39–77.
- [46] C.W. Hsu, C.C. Chang, C.J. Lin, A practical guide to support vector classification, Technical Report, 2003.
- [47] I.H. Jones, M. Pansa, Some nonverbal aspects of depression and schizophrenia occurring during the interview, *Journal of Nervous and Mental Disease* 167 (1979) 402–409.
- [48] J. Joshi, R. Goeck, G. Parker, M. Breakspear, Can Body Expressions Contribute to Automatic Depression Analysis?, in: *International Conference on Automatic Face and Gesture Recognition*.
- [49] M. Katsikitis, I. Pilowsky, A controlled quantitative study of facial expression in Parkinson’s disease and depression, *Journal of Nervous and Mental Disease* 179 (1991) 683–688.
- [50] D. Keltner, B.N. Buswell, Embarrassment: its distinct form and appeasement functions, *Psychological Bulletin* 122 (1997) 250.
- [51] G.L. Klerman, M.M. Weissman, B.J. Rounsaville, E. Chevron, *Interpersonal psychotherapy of depression*, Basic Books, New York, 1984.
- [52] E. Klinger, Consequences of Commitment to and Disengagement from Incentives, *Psychological Review* 82 (1975) 1–25.
- [53] B. Knutson, Facial expressions of emotion influence interpersonal trait inferences, *Journal of Nonverbal Behavior* 20 (1996) 165–182.
- [54] R. Kotov, W. Gamez, F. Schmidt, D. Watson, Linking big personality traits to anxiety, depressive, and substance use disorders: A meta-analysis, *Psychological Bulletin* 136 (2010) 768.
- [55] M.H. Mahoor, S. Cadavid, D.S. Messinger, J.F. Cohn, A framework for automated measurement of the intensity of non-posed Facial Action Units, in: *Computer Vision and Pattern Recognition Workshops*, Miami, pp. 74–80.
- [56] C.D. Mathers, D. Loncar, Projections of global mortality and burden of disease from 2002 to 2030, *PLoS medicine* 3 (2006) e442.
- [57] I. Matthews, S. Baker, Active Appearance Models Revisited, *International Journal of Computer Vision* 60 (2004) 135–164.
- [58] I.B. Mauss, R.W. Levenson, L. McCarter, F.H. Wilhelm, J.J. Gross, The tie that binds? Coherence among emotion experience, behavior, and physiology, *Emotion* 5 (2005) 175–90.
- [59] R. Mergl, P. Mavrogiorgou, U. Hegerl, G. Juckel, Kinematical analysis of emotionally induced facial expressions: a novel tool to investigate hypomimia in patients suffering from depression, *Journal of Neurology, Neurosurgery, and Psychiatry* 76 (2005) 138–140.
- [60] R.M. Nesse, Is Depression an Adaptation?, *Archives of General Psychiatry* 57 (2000) 14–20.
- [61] L.I. Reed, M.A. Sayette, J.F. Cohn, Impact of depression on response to comedy: a dynamic facial coding analysis, *Journal of Abnormal Psychology* 116 (2007) 804–809.
- [62] B. Renneberg, K. Heyn, R. Gebhard, S. Bachmann, Facial expression of emotions in borderline personality disorder and depression, *Journal of Behavior Therapy and Experimental Psychiatry* 36 (2005) 183–196.
- [63] J. Rottenberg, J.J. Gross, I.H. Gotlib, Emotion context insensitivity in major depressive disorder, *Journal of Abnormal Psychology* 114 (2005) 627–639.
- [64] J. Rottenberg, K.L. Kasch, J.J. Gross, I.H. Gotlib, Sadness and amusement reactivity differentially predict concurrent and prospective functioning in major depressive disorder, *Emotion* 2 (2002) 135–146.
- [65] A.J.J. Rush, M.B. First, D. Blacker (Eds.), *Handbook of Psychiatric Measures*, 2nd ed., American Psychiatric Publishing, Inc., Washington, D.C., 2008.
- [66] S. Sakamoto, K. Nameta, T. Kawasaki, K. Yamashita, A. Shimizu, Polygraphic evaluation of laughing and smiling in schizophrenic and depressive patients, *Perceptual and Motor Skills* 85 (1997) 1291–1302.
- [67] J.T. Schelde, Major depression: behavioral markers of depression and recovery., *The Journal of nervous and mental disease* 186 (1998) 133–40.
- [68] S. Scherer, G. Stratou, M. Mahmoud, J. Boberg, J. Gratch, A.S. Rizzo, L.p. Morency, Automatic Behavior Descriptors for Psychological Disorder Analysis, in: *International Conference on Automatic Face and Gesture Recognition*.
- [69] F. Schneider, H. Heimann, W. Himer, D. Huss, R. Mattes, B. Adam, Computer-based analysis of facial action in schizophrenic and depressed patients, *European Archives of Psychiatry and Clinical Neuroscience* 240 (1990) 67–76.
- [70] C. Segrin, Social skills deficits associated with depression, *Clinical psychology review* 20 (2000) 379–403.
- [71] C. Segrin, J.P. Dillard, The interactional theory of depression: A meta-analysis of the research literature, *Journal of Social and Clinical Psychology* 11 (1992) 43–70.
- [72] S.S. Shapiro, M.B. Wilk, An Analysis of Variance Test for Normality (Complete Samples), *Biometrika* 52 (1965) 591–611.
- [73] P.E. Shrout, J.L. Fleiss, Intraclass correlations: uses in assessing rater reliability, *Psychological Bulletin* 86 (1979) 420.
- [74] D.M. Sloan, M.E. Strauss, S.W. Quirk, M. Sajatovic, Subjective and expressive emotional responses in depression, *Journal*

- of Affective Disorders 46 (1997) 135–141.
- [75] D.M. Sloan, M.E. Strauss, K.L. Wisner, Diminished response to pleasant stimuli by depressed women, *Journal of Abnormal Psychology* 110 (2001) 488–493.
 - [76] L. Sloman, Evolved mechanisms in depression: the role and interaction of attachment and social rank in depression, *Journal of Affective Disorders* 74 (2003) 107–121.
 - [77] Y.I. Tian, T. Kanade, J.F. Cohn, Facial expression analysis, in: *Handbook of Face Recognition*, 2005, pp. 247–275.
 - [78] F. Tremeau, D. Malaspina, F. Duval, H. Correa, M. Hager-Budny, L. Coin-Bariou, J.P. Macher, J.M. Gorman, Facial expressiveness in patients with schizophrenia compared to depressed patients and nonpatient comparison subjects, *American Journal of Psychiatry* 162 (2005) 92–101.
 - [79] A. Troisi, A. Moles, Gender differences in depression: an ethological study of nonverbal behavior during interviews, *Journal of Psychiatric Research* 33 (1999) 243–50.
 - [80] J.L. Tsai, N. Pole, R.W. Levenson, R.F. Munoz, The effects of depression on the emotional responses of Spanish-speaking Latinas, *Cultural Diversity and Ethnic Minority Psychology* 9 (2003) 49–63.
 - [81] J. Whitehill, G. Littlewort, I.R. Fasel, M.S. Bartlett, J.R. Movellan, Toward Practical Smile Detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 31 (2009) 2106–2111.
 - [82] F. Wilcoxon, Individual Comparisons by Ranking Methods Frank Wilcoxon, *Biometrics Bulletin* 1 (1945) 80–83.
 - [83] M.A. Youngren, P.M. Lewinsohn, The functional relation between depression and problematic interpersonal behavior, *Journal of Abnormal Psychology* 89 (1980) 333.
 - [84] Z. Zeng, M. Pantic, G.I. Roisman, T.S. Huang, A survey of affect recognition methods: audio, visual, and spontaneous expressions, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 31 (2009) 39–58.
 - [85] Y. Zhu, F. De La Torre, J.F. Cohn, Y.J. Zhang, Dynamic cascades with bidirectional bootstrapping for spontaneous facial action unit detection, in: *International Conference on Affective Computing and Intelligent Interaction and Workshops*, pp. 1–8.



Figure 1: Example images of AU 12, AU 14, AU 15, and AU 24

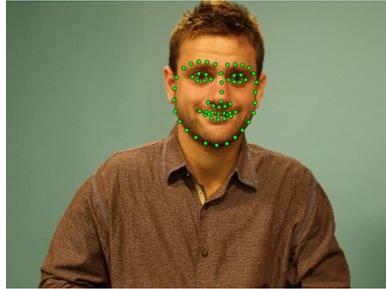


Figure 2: AAM Facial Landmark Tracking [19]

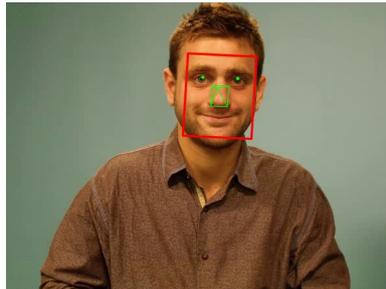


Figure 3: CSIRO Head Pose Tracking [20]

Table 1: Participant groups and demographic data

	Database	Responders
Number of Subjects	33	19
Number of Sessions	69	38
Average Age (years)	42.3	42.1
Gender (% female)	66.7%	63.2%
Ethnicity (% white)	87.9%	89.5%
Medicated (% SSRI)	45.5%	42.1%

Table 2: Reliability between manual and automatic FACS coding

Facial Expression	Frame-level	Session-level
AU 12	AUC = 0.82	ICC = 0.93
AU 14	AUC = 0.88	ICC = 0.88
AU 15	AUC = 0.78	ICC = 0.90
AU 24	AUC = 0.95	ICC = 0.94

Table 3: Facial expression during high and low severity interviews, Average base rate

	Manual Analysis		Automatic Analysis	
	High Severity	Low Severity	High Severity	Low Severity
AU 12	19.1%	39.2% *	22.3%	31.2%
AU 14	24.7%	13.9% *	27.8%	17.0% *
AU 15	05.9%	11.9% *	08.5%	16.5% *
AU 24	12.3%	14.2%	18.4%	16.9%

* $p < 0.05$ vs. high severity interview by Wilcoxon Signed Rank test

Table 4: Head motion during high and low severity interviews, Average root mean square

Head Motion	High Severity	Low Severity
Vertical Amplitude	0.0013	0.0029 *
Vertical Velocity	0.0001	0.0005 **
Horizontal Amplitude	0.0014	0.0034 **
Horizontal Velocity	0.0002	0.0005 **

* $p < 0.05$, ** $p < 0.01$ vs. high severity interviews by Wilcoxon Signed Rank test