HEBREW UNIVERSITY OF JERUSALEM

Automated Analysis of Non-verbal Behavior in Schizophrenia

by:
Talia TRON

Thesis for the degree
"Doctor of Philosophy"
in
Brain Sciences: Computation and Information Processing

Submitted to the Senate of the Hebrew University of Jerusalem

December 24, 2017

This work was carried out
under the supervision of
Prof. Daphna
WEINSHALL

To my family.

Acknowledgements

First and foremost, I would like to thank my advisor, Prof. Daphna Weinshall. Albert Einstein once said that *'It is the supreme art of the teacher to awaken joy in creative expression and knowledge.'*. If I need to describe our joint work over the last several years, the first words that come to mind is enjoyment. I will definitely miss our talks, our collaborative thinking, the exploration, the learning, and even our arguments. Whenever I was worn out from some aspects of the study, you gave me the motivation to look for a solution, explore new domains and find innovative perspectives. You encouraged me to carve my own path, and develop myself as a scientist and a researcher, while providing me with a safe environment where I can wonder, raise questions and make mistakes, knowing there is a 'responsible adult' looking over my shoulder. This has been a great journey and I'm privileged to have taken it with you, as an advisor, as a friend and as a colleague.

I would also like to thank my thesis committee - Prof. Abraham Peled and Dr. Oren Shriki who's comments and ideas were always enriching and invigorating. A special thanks to Prof. Peled, who's breakthrough thinking has made this dissertation possible. Thank you for inspiring me to take on the significant challenge of psychiatric diagnosis, and for consistently providing the patient's and doctor's perspective, keeping our work clinically relevant and substantial. I know we haven't touched upon all the problems you wish to solve in psychiatry, but the road is still ahead.

To the *Edmond and Lily Safra Center for Brain Sciences*, which was a second home for me in the past 6 years. My beloved classmates in the PhD program, thanks for making this journey so much fun, and the trip to Jerusalem always worthwhile. To Hezi, my office and research partner, it was great working with you, and I look forward to our next projects together. To ELSC's teachers, professors and staff, thanks for bringing us together from multiple disciplines and giving us the tools and knowledge needed to become top scientists. A special thanks to Prof. Eli Nelken, head of the PhD program who was like a (tough) father for me and my classmates during the first academic years and encouraged us to always find the elephant in the room.

Last but not least, I would like to thank my family, a place where creativity, integration and out-of-the-box thinking is the name of the game. My parents, you gave me the courage to always follow my heart, and the surrounding needed to succeed. My older sister and brothers, your bright advice, your experience and support, accompany me wherever I go. To my dear husband Itai, who consistently motivates me in his quiet, intelligent way. I wouldn't have become a doctor without you. Finally, to my son, Eyal, the love of my life.

HEBREW UNIVERSITY OF JERUSALEM

Abstract

Edmond and Lily Safra Center for Brain Sciences

Doctor of Philosophy

Automated Analysis of Non-verbal Behavior in Schizophrenia

by Talia TRON

The descriptive non-quantitative manner in which psychiatric disorders are diagnosed, makes it hard to monitor patients, evaluate treatment influence, and compare between different symptomatology. In recent years, technological and algorithmic developments enable us to extract and analyze measures of nonverbal behavior such as facial expressiveness, body gestures and prosody, which are an integral part of psychiatric diagnosis.

The objective of this dissertation is to characterize nonverbal behavior in schizophrenia, and to develop automatic tools for quantitatively describing and analyzing relevant measures of this behavior. We focus on facial expressions and motor behavior, and combine descriptive statistical methods together with data-driven analysis techniques to obtain a wide range of nonverbal characteristic measures. These measures include the intensity, dynamics, consistency and appropriateness of facial and motor behavior. The obtained nonverbal measures are then used to refine the definition of nonverbal alterations in schizophrenia, explore the relation between them, and describe the way they are manifested in a clinical setting. In addition, these measures are used to automatically classify clinical sub-population, evaluate symptom severity and identify significant irregularities in patients behavior over time.

We hope that the methods and approaches for automatic non-verbal behavior analysis introduced here, will contribute to the field of affective computing in general, and psychiatric diagnosis in particular, will increase the reliability of the diagnostic procedure, and will allow better characterization and monitoring of patients' behavior, which will promote both research and treatment.

Letter of Contribution

Chapter 2.1 - Automated Facial Expressions Analysis in Schizophrenia: a Continuous Dynamic Approach

Chapter 2.2 - Facial Expressions and Flat Affect in Schizophrenia, Automatic Analysis from Depth Camera Data

Chapter 2.3 - Differentiating facial incongruity and flatness in schizophrenia, using structured light camera data.*

Talia Tron, Abraham Peled, Alexander Grinsphoon, and Daphna Weinshall.

TT, AP and DW, planned and carried out the experiment. AP Devised the main conceptual ideas and developed the theory conceived in the presented ideas, he supervised the experiment and aided in clinically interpreting the results. TT designed and directed the experiment, supervised data collection, performed the presented analysis, interpreted the results together with TT and DW, and wrote the majority of the manuscript. DW supervised and guided TT in results analysis, verified the analytic methods and tools, contributed to the interpretation of the results and to the writing of the manuscript. AG was involved in planning and supervising the experiment.

All authors discussed the results and contributed to the final manuscript.

*The above description holds for the 3 abovementioned papers.

Chapter 3.1 - Real-time Schizophrenia Monitoring using Wearable Motion Sensitive Devices

Talia Tron, Yehezkel S. Resheff, Mikhail Bazhmin2, Abraham Peled and Daphna Weinshall.

TT and YR devised the main conceptual ideas, planned and designed clinical experiment. TT has designed experimental interfaced, directed and supervised data collection, while YR has monitored data storing and preprocessing. MB collected the experimental data, and performed the clinical evaluations under the supervision and guidance of AP and TT. When all data was collected, YR derived the motor features (step count, energy etc.), and TT performed the correlations with clinical data over different time windows. Both authors, together with DW and AP, contribute to the interpretation of the results and writing of the manuscript. AP and MB gave the theoretical clinical knowledge for experimental design and interpretation.

Chapter 3.2 Topic Models for Automated Motor Analysis in Schizophrenia Patients

Talia Tron, Yehezkel S. Resheff, Mikhail Bazhmin, Abraham Peled, Alexander Grinshpoon and Daphna Weinshall.

Yehezkel S. Resheff and Talia Tron are equal contributors to this manuscript. They both devised the main conceptual ideas, planned and designed clinical

experiment. TT has designed experimental interfaced, directed and supervised data collection, while YR has monitored data storing and preprocessing. When all data was collected, YR derived motor 'code-book' and performed topic model analysis on the words generated. TT has processed the topic representation to obtain 'consistency', 'typicality' and 'richness' measures, analyzed clinical data to derive sub-clinical populations over time and designed a classifier for automatic clinical classification. Both authors, together with DW, contribute to the interpretation of the results and writing of the manuscript. AP AG and MB gave the theoretical clinical knowledge for experimental design and interpretation. MB collected the experimental data and performed the clinical evaluations under the supervision and guidance of AP and TT. All authors discussed the results and contributed to the final manuscript.

Chapter 3.3 ARIMA-based motor anomaly detection in schizophrenia inpatients

Talia Tron, Yehezkel S. Resheff, Mikhail Bazhmin, Abraham Peled, and Daphna Weinshall.

TT and YR devised the main conceptual ideas, planned and designed clinical experiment. TT has designed experimental interfaced, directed and supervised data collection, while YR has monitored data storing and preprocessing. MB collected the experimental data and performed the clinical evaluations under the supervision and guidance of AP and TT. When all data was collected, YR designed a step count detector using raw motor data. TT used the step count feature to design an ARIMA model for patients' abnormal behavior detection and systematically compared its performance with clinical and medication data. Both authors, together with DW and AP, contribute to the interpretation of the results and writing of the manuscript. AP and MB gave the theoretical clinical knowledge for experimental design and interpretation.

"The most important thing in communication is hearing what isn't said."

Peter F. Drucker

Contents

Chapter 1

Introduction

Psychiatric diagnosis and nonverbal behavior

In psychiatry, mental disorders are diagnosed in a descriptive non-etiological manner based on patients' complaints, medical condition, psychiatric history and on the evaluation of their mental status. Mental status examination (MSE) is done using lists of signs and symptoms, which include nonverbal observations such as facial and vocal expressions, since they convey extensive information about the patient's emotional and mental state [82, 21, 24]. Bodily expressions may be as informative as voice in clinical settings, and even as facial expressions, although they are hard to map into discrete emotional conditions [38].

The diagnostic process is based on the *Diagnostic and Statistical Manual of Mental Disorders* (DSM), a handbook written and published by the *American Psychiatric Association* (APA), which contains prototype categories for more than 200 mental disorders [8]. The DSM was criticized for having low inter-rater reliability, for lacking empirical evidence and for being subjected to personal interpretation. The fluid, non-evidence based nature of the DSM makes it vulnerable to political and economical interests rather than clinical ones [30, 83, 20]. Under the DSM framework, nonverbal behaviors are non-accurately described using general terms such as 'motor retardation' and 'flat affect'. More specific clinical scales, such as the *Positive and Negative Syndrome Scale for Schizophrenia* (PANSS) [53] and the *Hamilton Rating Scale for Depression* (HRSD) [40] contain items which directly address patients nonverbal behavioral symptoms. Nonetheless, to date, these symptoms are evaluated based on the subjective impression of the diagnosing psychiatrist which may be biased by their life experience, ideation and culture. In addition, symptom evaluation process requires expert staff and availability of resources, and it is not done frequently enough to capture delicate changes in patients' spontaneous and drug-induced condition [103]. To date, there are no objective,

quantitative methods to measure nonverbal behavior in psychiatric patients, which causes multiple interpretations of phenomenology and results in low reliability and validity of psychiatric diagnosis [21].

In 2008 the *American National Institute of Mental Health* (NIMH) have revolutionized the field of mental disorder research by announcing that studies which use the DSM as a clinical criteria will no longer be supported and funded. Instead, they began implementing a different approach, manifested in the *Research Domain Criteria* (RDoC) framework. The RDoC framework shifts the focus from high level categories such as 'schizophrenia' and 'depression', to more refined subcategories homogeneous across different mental disorders (e.g. attention, visual perception and facial communication). The rational is that bottom up data accumulation will create the evidence foundation needed for integrating circuit, neural network and behavioral levels, and will allow better understanding of the underlying mechanisms of psychiatric conditions [45]. The work presented in this dissertation may contribute to the RDoC effort by providing tools for objective and detailed description of nonverbal phenomenology and its relation to different behavioral and clinical aspects.

Non verbal behavior in schizophrenia

We focus on schizophrenia, one of the most severe mental disorders with lifetime prevalence of about 1% worldwide. The disorder is characterized by *negative symptoms*, which involve the loss of functions and abilities (e.g. lack of speech and motivation, blunted affect), and by *positive symptoms*, which are pathological functions not present in healthy individuals (e.g. auditory hallucinations, delusions and paranoid thoughts).

Psychiatrists generally distinguish between two types of schizophrenia. Type I are psychotic disorganized patients characterized mainly by positive symptoms (*positive-signs schizophrenia*). In clinical setting, these patients show disorganized speech, restless behavior and lack of consistency. Their facial expressiveness tend to be less suitable to their subjective feeling and to the external situation, and their motor behavior is characterized mainly by irritability, involuntary movements, dyskinesia and catatonic symptoms [80, 106, 78, 108, 102]. Type II are patients who demonstrate post-psychotic residual *negative-signs schizophrenia*. These patients nonverbal behavior is characterized by reduced facial expressiveness, monotonic speech and observed slowness and psycho motor poverty [73, 72, 18, 60, 27, 107]. Some

patients demonstrate both types simultaneously or during different phases of the illness. Treatment is mainly through neuroleptic drugs which relieve the positive symptoms but are less effective against the negative ones [7]. The neuroleptic treatment may also result in drug-induced movement disorders such as tremor dystonia, parkinsonism (rigidity and bradykinesia), akathisia and tardive dyskinesia, mainly in chronic patients [10, 47, 59].

The diversity and specificity of facial alterations and motor symptoms throughout different phases of schizophrenia disorder and as a response to drugs, makes them good candidates for patients monitoring and treatment outcome evaluation. Nonetheless, as is the case in other aspects of psychiatric diagnosis, these are evaluated subjectively and qualitatively with low inter-rater reliability. Consequently, it is useful, and often time critical, to develop an assistive technology that can achieve more objective nonverbal activity assessments, and further refine the often-subtle discrimination between motor and different affect-related symptoms. Such automatic nonverbal measures may allow for a rich, quantitative and continuous monitoring of patients' behavior over time, making it easier to detect delicate changes in their clinical condition and compare clinical phenotypes between patients.

Automated nonverbal analysis tools

In the two past decades there was a considerable effort to automatically extract and analyze human nonverbal behaviors such as facial expressions, vocal prosody and bodily gestures. The extraction of nonverbal behavioral features has improved vastly in recent years, thanks to some technological and algorithmic innovations including 3D cameras for facial expressions and embedded actymetric sensors for motor activity (see Sections 1.1.2 and 1.2.1 respectively). The field has become known as *Affective Computing* (AC) or *Social Signal Processing* (SSP). AC studies mainly focus on the emotional aspect of nonverbal behavior, while SSP studies gives greater weight to interpersonal communication and takes into account measures like context and environment [101, 99, 38].

An emerging field of study focuses on implementing AC/SSP techniques to medical data in order to improve diagnosis, monitoring and treatment. In such studies, subject's behavior (body, voice or face) is recorded during an interview, nonverbal behavioral features are extracted manually or automatically, analyzed, and correlated with medical condition or evaluated by

clinical tests. The growing body of work in the field includes different neurological and psychiatric states such as ADHD [75, 52, 81], autism [77, 110] depression [99, 34, 55, 35] and some studies about schizophrenia [41, 42, 57, 62, 71, 93], which will be further discussed in Sections 1.1.4 and 1.2.3.

Most clinical studies take a *descriptive* approach, where basic behavioral measurements or combination of measurements are given a semantic meaning by the researcher (for example 'anxiety level', 'happiness level'). Such a semantic interpretation may be compelling, but should be done with great caution, since it relies heavily on theoretical assumptions which in some cases are still under debate. A good example is the continuous debate between prominent researchers in the field of psychology and sociology over the emotional interpretation of facial activity data (see Section 1.1.1).

To avoid possibly misleading theoretical assumptions, some studies address behavior directly, without looking for the message underlying it, taking a *data-driven* approach. In such studies classic statistical methods are combined with machine learning tools to derive nonverbal behavioral measures of various kinds. These are then often used for automated clinical classification, prediction of patients symptom severity or digitized evaluation of treatment outcome. The main advantage of such methods is the objective manner in which measures are derived and analyzed, reducing possible statistical and conceptual biases. Notwithstanding, data-driven methods typically require great amount of labeled data, and researchers in the field tend to use broad, general measures of nonverbal behavior which can hardly be construed as representing human behavior (e.g. general movement measurements commonly used in computer vision). There is a fine line between over interpreting behavioral features and using unintelligible measures, non-relevant for clinical diagnosis.

This dissertation can be seen as an implementation of AC techniques for the study of schizophrenia, with a focus on facial expressions and motor behavior. We attempt to define data-based ecologically relevant measures, which will be accurate and informative enough to be used as input for learning and prediction algorithms, and self-explanatory and ecologically relevant enough to gain clinical insights.

Research objectives

The main objectives of this dissertation are: (1) to characterize nonverbal behavior in schizophrenia patients, focusing on facial and motor activity; and

(2) to develop automatic tools for quantitatively describing and analyzing relevant measures of this activity. In order to achieve these objectives, we combine descriptive statistical methods together with data-driven analysis techniques including time series forecasting, machine learning and natural language processing (NLP) algorithms. With these techniques, we obtain a wide range of nonverbal characteristic measures. These are then used to refine the definition of nonverbal alterations in schizophrenia, explore the relation between them, and describe the way they are expressed in a clinical setting. In addition, these nonverbal measures are used to automatically classify clinical sub-population, evaluate symptom severity and identify significant irregularities in patients behavior over time.

Our hope is that automatic extraction and analysis of such behavioral measures will contribute to the reliability of psychiatric diagnosis, and allow better characterization of patients' behavior, which will promote both research and treatment.

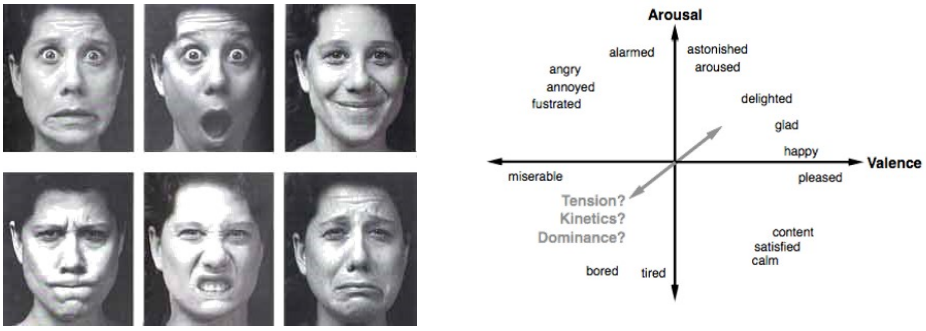
In the remainder of the introduction we elaborate the discussion separately on facial expressions (Section 1.1) and motor analysis (Section 1.2). We will start with an overview of prominent technological and theoretical approaches in each field, followed by a literature review of AC/SSP studies done in schizophrenia, and conclude with a brief description of our work and its main innovations.

1.1 Facial Expressions

1.1.1 Theoretical approaches

The face has a major role in signaling people's emotional and mental state. Subsequently, most traditional studies in psychology and sociology as well as AC/SSP studies focused on facial expressions and their relation to emotions [99, 38, 97]. The work in the field may be roughly divided into the *Categorical approach* vs. the *Dimensional approach* (Figure 1.1). The *categorical approach* was pioneered by Paul Ekman, who in a set of extensive studies done since the 50's pointed out 6 prototype expressions universally expressed and recognized - anger, fear, sadness, happiness, disgust and surprise [22]. Ekman's critics argue that he uses exaggerated, static and posed facial expressions, while those presented in everyday life are dynamic, spontaneous and far more subtle [38, 9]. In contrast, the *dimensional approach* characterizes emotions according to two or more axes, usually including valence (negative to

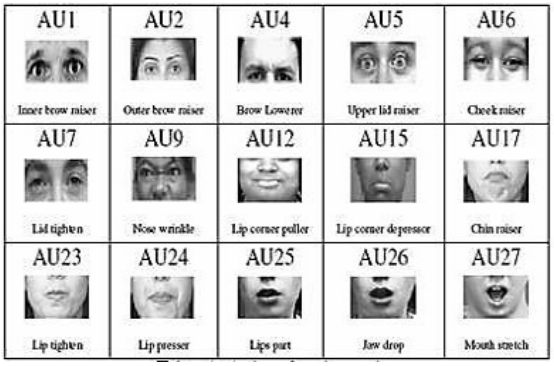
positive feelings) and arousal [84, 86]. These dimensions may capture greater variety and subtleties of human emotions, but are harder to report and measure. Both approaches focus on the *emotional* aspect of facial activity while neglecting other characteristics such as the amount of variability and the way expressions change over time. Overall, the vast majority of studies in the field are limited to categorical, static and posed expressions while modeling natural behavior in a continuous, dynamic manner remains a challenge only few research take [48].



(A) (B)

FIGURE 1.1: Illustration of the two most prominent approaches for emotional modeling. (A) The categorical approach which maps facial expressions to 6 prototype emotions (anger, fear, sadness, happiness, disgust and surprise), (B) The dimensional approach which describe emotions on valence and arousal axes.

An alternative approach is to analyze the facial activity without interpreting its emotional state. This is commonly done using the *Facial Action Coding System* (FACS), originally developed by the Swedish anatomist Carl-Herman Hjortsjö [43]. FACS gives a score to the activity of roughly 46 individual facial muscles called *Action Units* (AUs) based on their intensity level and temporal segments (*Figure 1.2-A*). Neuroscientists, psychologists and sociologist have been using FACS for the past 30 years to conduct extensive research regarding various aspects of facial expressions. In the early 80's, Ekman adopted FACS and developed the *Emotional Facial Action Coding System* (EMFACS) which systematically categorizes combination of AUs to specific emotions [31] (Figure 1.2-B). This allowed those in favor of the *categorical approach* to implement FACS in their studies. To date, there is no systematic transformation from FACS coding to emotional dimensions of valence and arousal. In a parallel effort, we addressed this problem using emotionally evocative *YouTube* videos, with some reassuring results [39].



(A)

Emotion	Action Units
Anger	4+5+7+23
Contempt	R12A+R14A
Disgust	9+15+16
Fear	1+2+4+5+7+20+26
Happiness	6+12
Sadness	1+4+15
Surprise	1+2+5B+26

(B)

FIGURE 1.2: Illustration of (A) the Facial Action Coding System (FACS), which gives a score to the activity level of of independent facial muscles (facial Action Units), and (B) the Emotional Facial Action Coding System (EMFACS), which systematically categorizes combination of facial Action Units to prototype emotional expressions.

1.1.2 Facial features extraction tools and technology

While automatic prototype expression detection may be considered by some a solved problem [51], automatic FACS coding that will replace the manual one still poses a major challenge in the field of machine learning and computer vision [76]. The problem is of high complexity given the large number of classes (AUs), the subtle, small different between them, and the lack of standardized, large quantities labeled data-sets. Subsequently, deep learning methods, which have dramatically enhanced performance in other computer vision challenges, have only been moderately beneficial for AUs recognition as yet [67, 37, 46].

Leaving deep learning methods aside, automatic AUs extraction may be done using *Geometric-Features based* methods, based on information regarding location of points or shapes on the face (e.g. position, speed, acceleration) [66, 63] or using *Appearance-Based* methods based on changes in texture and motion of the skin such as wrinkles and furrows [61, 12, 5]. The use of geometric features usually yields better results, since appearance based methods are more sensitive to illumination conditions and to individual differences, though a combination of both methods may be preferable [96]. A newer, promising method is based on temporal information in AU activity, which was found to improve recognition as compared to static methods [68, 56].

Basic features are classified into AUs using model driven methods such as active appearance models (AAMs) [48, 64, 98] or data driven methods [89].

Data-driven methods require larger sets of data in order to cope with variations in pose, lighting and textures, and they allow a more accurate person independent analysis of AUs expression [91].

3D cameras based on structured light technology enable the capture of facial surface data which is less sensitive to head pose and to lighting conditions as compared with 2D data. These cameras project infra-red patterns onto the 3D scene, and measure the deformation created by objects, yielding a better recognition rate of AUs [87]. A drawback, however, is that the depth resolution is rather low, and that the image may contain small artifacts in highly reflective and non-reflective regions, or holes in regions not covered by the projector [90].

The most prominent developments in the field have become commercialized and are available as software for industry and academic use [25, 44, 1].

1.1.3 Facial Expressions in schizophrenia

Facial expressions are essential to the definition and description of schizophrenia. Blueler, the swiss psychiatrist and eugenicist who coined the term at the beginning of the last century, described *Affective incongruence* as a main characteristic of the disorder. He clinically defined it as *'a discrepancy between the affect and the content of speech, subjective feeling, or the situation'* [17]. This discrepancy is typically manifested in the early psychotic stage of the disorder *(positive-signs schizophrenia*). As the disorder progresses to deficiency *negative-signs schizophrenia*, affect usually becomes more constricted, with a severe reduction in emotional expressiveness clinically known as *Affective Flatness* [69]. There is evidence for high congruence between flat affect severity, patients' wellbeing and treatment outcome [3].

Traditionally, *Affective incongruence* is thought to be independent of *Affective Flatness* and is therefore measured separately using different clinical and empirical scales [65]. Nonetheless, the reports of incongruence may sometimes merely be the result of misinterpretation of facial flattening, depending on the context. For example, a 'frozen smile' which does not change over time, may be considered as facial flattening, but when the patient is asked to describe how he or she feels, it can be interpreted as inappropriate. In this dissertation we made an effort to overcome this confound, and to propose a way to distinguish between flat and incongruent affects (See Figure 1.3).

Another possible clinical bias may arise in the emotional interpretation of facial activity. The *Inhibition Theory* states that the emotional experience

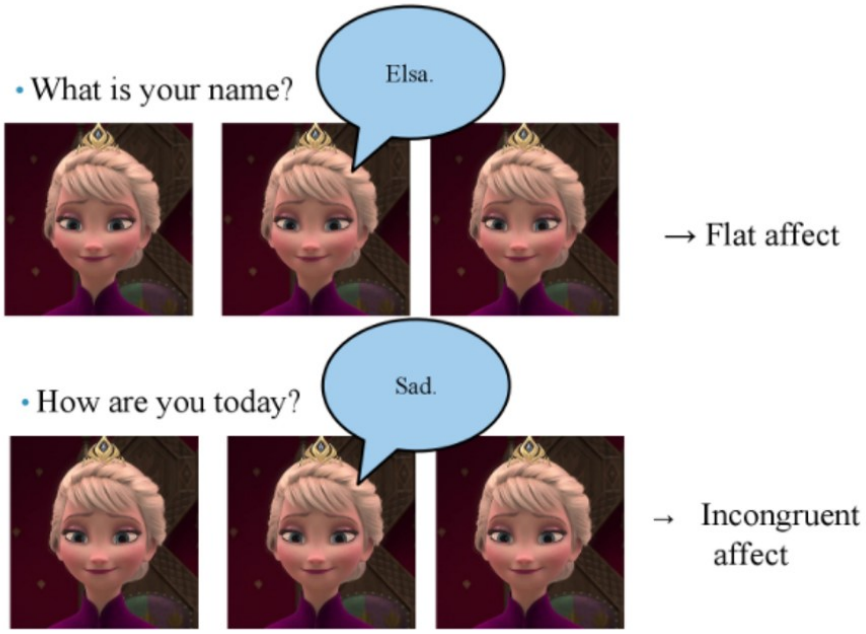


FIGURE 1.3: Illustration of the possible confound between *Affective flattening* and *Affective incongruence* where the same facial behavior (e.g. a frozen smile) can be interpreted as both, depending on the behavioral context.

is not compromised in schizophrenia, but rather that there is an impairment in emotional expression [74, 2]. In fact, Blueler himself interpreted incongruence as a 'split' between mental mechanisms controlling affect and those controlling mood experiences, resulting in a discrepancy between patients' mood and their affect. Subsequently, any emotional interpretation of patients behavior based on their facial expressions should be done with maximum caution and sensitivity, as we will demonstrate in our study.

1.1.4 Previous work

Only a minority of AC/SSP studies dealt with schizophrenia. These studies may be roughly divided into the ones taking descriptive approach (interpreting the measures), versus the ones using data-driven analysis, with no emotional or mental interpretation (both approaches are discussed in detail in Section 1).

The descriptive studies focus mainly on the *affective* aspect of patients' nonverbal behavior, namely, on the emotional value of their facial expressions. These studies commonly use the *Emotional Facial Action Coding System*

(EMFACS) to interpret AUs activity in terms of basic emotions such as happiness, anger or fear. Based on this interpretation, studies have found patients to demonstrate less positive emotions than controls [62], that the dominant emotion during an interview is disgust [28] or contempt [95, 92], and that the congruity of emotional response is reduced in patients [16]. The drawback of such studies however, is that they heavily rely on the *Categorical approach* for emotional interpretation (described in Section 1.1.1), and do not allow for a more subtle, real-life characterization of patients facial response. In addition, they do not necessarily take into account the possible discrepancy between emotional experience and emotional expression (see the *Inhibition Theory* described in previous section), making their conclusions somewhat superficial.

Another measure often used in studies is facial flatness, usually calculated as the intensity of specific AUs, or overall AUs. Results indicate reduced upper facial activity [23] and reduced overall facial expressiveness in patients in reaction to emotional stimuli [16, 93, 26] and to social interaction [6, 32]. In a recent study by Vijay et al. [100, 11], smiling behavior was found to be negatively correlated with negative symptoms severity. In addition, unusual thought content and delusions were associated with more brows raise.

The aforementioned studies used a limited set of nonverbal characteristic features and ignored information regarding facial dynamics and variability. Furthermore, none of them used data-driven methods to obtain a more comprehensive description of patients' behavior. The few data-driven studies done on facial expression in *depression*, though outpointing some interesting results, used ecologically irrelevant measures (e.g. general measurements of motion, commonly used in computer vision) and a broad, non-reliable psychiatric definition for data labeling ('level of depression') [50, 49].

Despite the crucial part of *affective incongruent* for schizophrenia diagnosis, only one study so far has derived facial congruity measures in schizophrenia patients. This was done in a study by Hamm et al., who analyzed facial activity under different emotional conditions using information theory measures. They found patients to be lower in terms of *distinctiveness* (mutual information between AU activity and emotional condition) and higher in *ambiguity* (entropy of AU activity during a specific emotional condition) [42]. Nonetheless, this study did not tackled the possible confound between flatness and congruity, limiting the clinical importance and relevance of their results.

1.1.5 Current work

In this dissertation we used a structured light depth camera and dedicated software to characterize spontaneous facial activity in schizophrenia patients. Based on this characterization, we targeted the clinical definitions of *Affective flatness* and *Affective incongruity*, suggesting a more subtle, detailed, observation-based definition. We further investigated how these symptoms can be measured, and preferably distinguished, by automatic means from recordings of spontaneous facial expressions of schizophrenia patients. Finally, we investigated the possibility of accurately detecting and quantifying affect as effectively as clinicians do in their regular mental status examination.

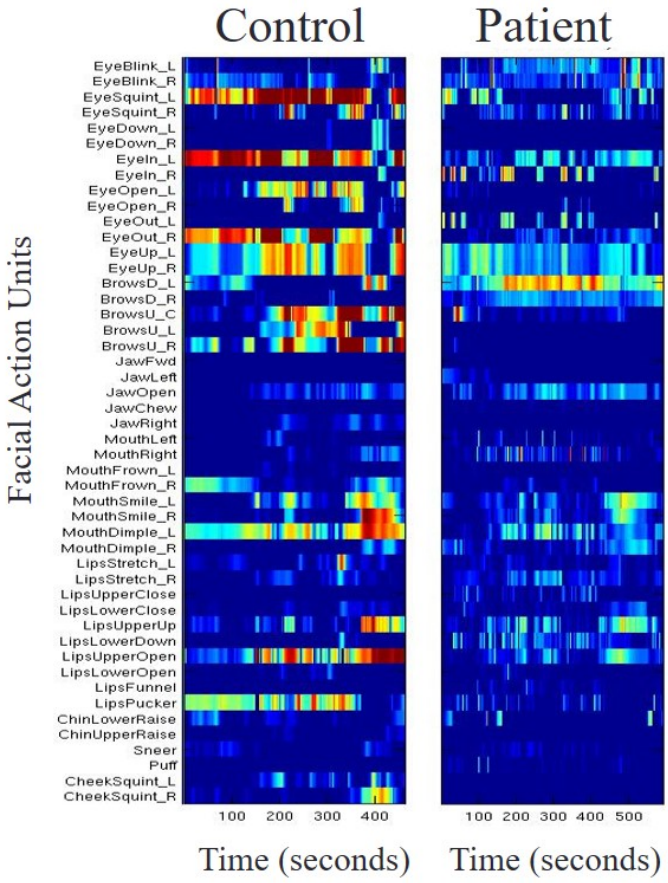


FIGURE 1.4: Illustration of raw facial data for one patient and one control subject during short structured interview. Each line represent a different facial Action Unit (AU), with warmer colors representing higher level of activity.

In the first two papers we compare the facial activity of schizophrenia patients and healthy individuals during a short, structured interview. We expand the definition of *Affective flatness* from mere facial intensity (how strong are facial expressions in terms of muscle activity?) to facial dynamics (how

much do they change over time?) and richness (how much variability is there? Namely, how many different expressions?).

In paper 2.1, looking at each facial AU separately, we characterize the facial intensity and dynamics of patients using transition matrix representation, and present a two-step SVM based algorithm designed for patients vs. control classification and symptom severity evaluation.

In paper 2.2, we obtain a dense representation of data-driven 'prototype' expressions using clustering analysis over all facial AUs (illustrated in Figure 1.5). We then suggest to compartmentalize the definition of *Affective flatness* to the following three components - **Richness** (how many expressions appeared?), **Typicality** (how similar they were to the prototype?), and **Affective Distribution** (which expressions were more prevalent?). These components are then compared between patients and controls, and finally correlated with clinical flat affect severity assessed by the psychiatrist.

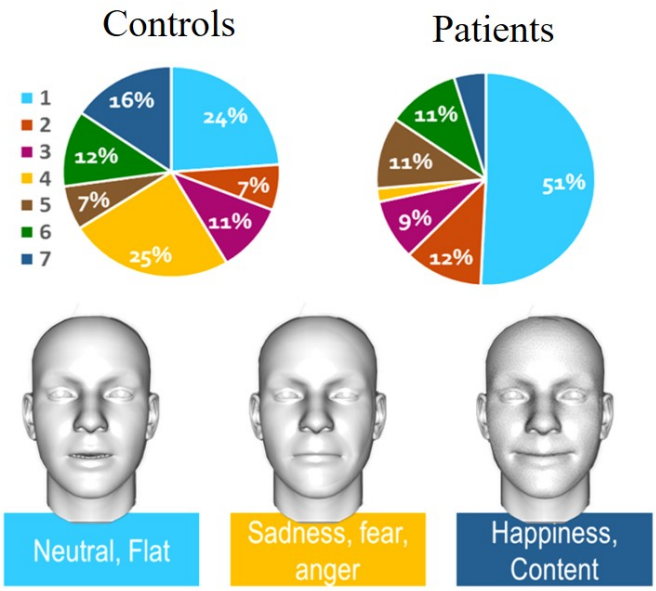


FIGURE 1.5: Illustration of three facial clusters, their mean distribution in patients and control subjects, and their emotional interpretation based on the Emotional Facial Action Coding System (EMFACS). While facial activity of control subjects is distributed over all different facial expressions (clusters), patients demonstrate mainly neutral or flat expression.

In paper 2.3, we address the definition of *Affective Incongruence*, aiming to differentiate inappropriateness from *Affective flatness*, and rule out possible emotional deficits. Our methods were as follows: First, we asked patients

and control subjects to rate their emotional response to emotionally evocative photographs, and subsequently compared their ratings (**Emotional Congruity** score). Then, we conducted variance analysis and reformulated the term facial incongruity to consistency (how similar was facial response to his own response when watching similar emotional stimuli?) and appropriateness (how similar was the facial response the typical expression in the healthy population?). Importantly, in order to measure whether the level of incongruity goes beyond mere facial flatness, we normalized both scores by the overall flatness.

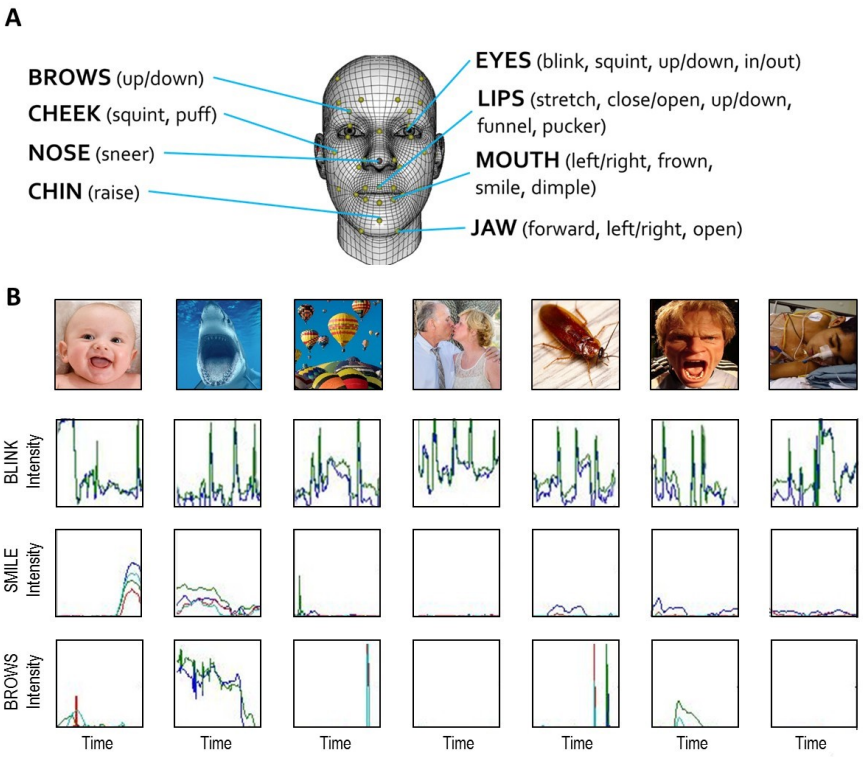


FIGURE 1.6: a) Illustration of the facial Action Units (AUs) automatically tracked in this study using 3D video and dedicated software. b) Raw AUs activity of one subject while watching emotionally evocative pictures retrieved from the International Affective Picture System (IAPS)

To our knowledge, this is the first attempt to use automated analysis tools in order to tackle clinical definitions of affect in schizophrenia. Likewise, no study so far has empirically distinguished affect incongruence from mere flatness. In this sense, our study offers a new perspective on facial behavior in patients and demonstrate the huge potential of automated affect evaluation as an assistive tool in clinical settings.

1.2 Motor Behavior Analysis

1.2.1 Monitoring with wearable devices

Actigraphy, namely, non-invasive monitoring of human motor activity, often via wearable sensors (also known as *actigraphy*), has become widely available in recent years, thanks to some technological developments making it cheaper, more accessible and easy to use. The most prevalent actimetry sensors are accelerometers, which measure the change in velocity in 3 dimensions, and gyroscopes which measure orientation and angular velocity. In healthy individuals actimetry sensors are commonly used to measure sleep quality [88], physical activity (wake-time activity, calories burned estimation) and other movement measures such as the number of steps per period (step count), the overall energy, and the overall dynamic body acceleration (ODBA) [13]. In addition, physiological sensors for heart rate and skin conductance combined with accelerometers enable the detection of high heart rate and arousal which are not associated with physical activity [4].

Wearable actimetry devices are worn on the body (usually wrist or chest) or otherwise attached or embedded in it. Actigraphy is also commonly monitored using mobile devices (smart phones), in combination with additional behavioral measurements such as location (GPS), screen use, and anonymized call and text message logs. The advantages of using mobile devices include their high availability in clinical and healthy population, and the ability to interact with users and get feedback. Nonetheless, theses devices are not carried continuously on users body, making them less sensitive to subtle motor changes, and subsequently their collected data is far less accurate and reliable [29]. Subsequently, in this dissertation we choose to focus on accelerometer sensors embedded in smart watches.

The last decade has seen a steep rise in use of wearable actimetry devices in medical fields ranging from human physiology [94] to movement disorders [105, 58] and mental health [104]. In a clinical setting, these devices may be used in order to detect change in high-level movement parameters, track their dynamics and correlate them with various behavioral and clinical parameters. In Parkinson's disease research, measures of mobility [105] as well as measures of specific motor related symptoms [54] were developed to aid monitoring and personalization both in clinical and home settings. In the mental health field, a platform was developed for collection of heart-rate and accelerometer data, as well as user self-report questionnaires [33], in order to better understand the factors associated with mental health and well-being in

students. The PSYCHE system [79] was designed to collect data from mood disorder patients, using textile platforms and other portable sensors, in order to monitor patients, and facilitate interaction between patient and physician. The system is also capable of alerting the medical staff when a manic episode is detected.

Recent developments together with empirical findings increase the feasibility of using such sensors to monitor schizophrenia patients in natural and clinical settings. Studies which tested schizophrenia patients' compliance to wearable devices, found that they tend to cooperate and do not report any significant discomfort from begin recorded [36]. Mobile behavioral sensing using smart phones was also demonstrated to be acceptable and informative for data collection in outpatients and inpatients with schizophrenia [14].

Notwithstanding the above, to date there have been no attempts at a complete, self-contained system, to directly assess and monitor schizophrenia patients using wearable devices and sensors. Such a system could have great potential due to the illusive nature of the disease, and the overwhelming need for sub-typing towards a better understanding of underlying causes, and development of better and more personalized treatment.

1.2.2 Motor behavior in schizophrenia

Motor behavior is an integral part of schizophrenia disorder and is essential for the diagnosis and evaluation process. Different phases of the disorder are characterized by unique patterns of motor alterations. Patients with *positive-signs schizophrenia* occasionally show severe motor deficits, with increased overall motor activity, irritability, increased involuntary movement, and decreased voluntary movements (dyskinesia). Catatonic symptoms are also manifested in this phase, expressed as constant hyperactivity, or rigid motor poses indifferent to external stimuli [80, 106, 78, 108, 102]. In *negative-signs schizophrenia* there is usually an overall retardation of motor activity, characterized by slowness, decreased spontaneous movements, and psychomotor poverty [73, 72, 18, 60, 27, 107]. The reduced facial activity observed in patients (flat affect) can also be seen as part of the negative-signs motor deficits, as discussed in Section 1.1.3.

People with schizophrenia tend to do less Physical Activity (PA), making them more vulnerable to coronary heart disease and metabolic syndromes [109]. In addition, their sleep quality is often disturbed, which was shown to be related with symptom severity, psychotic relapse, premature mortality,

and even suicide [85]. In accordance, direct continuous monitoring and real time interventions to improve PA levels and sleep quality for schizophrenia patients, may help to improve their reduced life expectancy.

The effect of medical treatment on motor symptoms of schizophrenia is double-sided. On the one hand, the medical treatment may improve the neurological symptom signs (NSS) present in the early phases of the disorder, reduce the level of involuntary movement and decrease the amount of dyskinesia [80, 19]. On the other hand, the medications may cause motor adverse effects, mainly in chronic patients, including tremor, repetitive or sustained muscle contractions (dystonia), parkinsonism (slowness, stiffness and speech impairment) and akathisia (feeling of motor restlessness with the need to be in constant motion) [10, 47, 59].

1.2.3 Previous work

Only a handful of studies in the field of wearable devices has dealt with schizophrenia. Wichniak et al. [109] used actigraphy to test daily and 24-hours motor activity in schizophrenia patients (n=73), and found both to be reduced in comparison to control subjects (n=36), where in addition patients spent more time in bed. Lower motor activity was found to be correlated with negative symptoms and depressive symptoms. Similar findings were described by Yamamoto et al. [70], who measured the amount of daily physical activity (PA) in schizophrenia (n=37) and control group (n=41), revealing lower intensity level and shorter PA duration in the schizophrenia group. Here also, reduced activity was correlated with negative symptoms. Despite their obvious importance, these studies use very general motor activity descriptors, with a main focus on the motor intensity and amplitude measures, ignoring other parameters which might be relevant (as is the case for facial expressions).

A different approach was suggested in a study by Berle et al. [15] who recorded two weeks of actigraphy data in schizophrenia and major depression patients. Patients' motor behavior was described in terms of three variables: the inter-daily stability (IS), intra-daily variability (IV), and the relative amplitude (RA). In accordance with other studies, the relative amplitude of motor activity in both schizophrenic and depressed patients was found to be significantly reduced. In addition, schizophrenia patients treated with clozapine demonstrated a more structured behavioral pattern, expressed in higher intra-daily stability (IS), and lower intra-daily variability (IV). In a

recent study, Reinertsen et al. [85] trained a SVM classifier to distinguish between schizophrenia and control group (n=12 in each group) based on daily heart rate and PA measures. The classifier used classical statistical features such as rest-activity metrics and transfer entropy (a measure of activity disorganization over time) and resulted in an area under the receiver operating characteristic curve (AUC) of 0.96. The algorithm was evaluated on different time windows, and the results demonstrate how daily estimation of illness severity may be done using continuous patients monitoring over short time scales. These studies demonstrate the potential of using temporal information to gain more subtle measures of motor variability and stability.

All of the above studies suffer from several problems. First, their analysis is restricted to group comparison and symptom severity evaluation. No effort has been made to define and characterize different clinical sub-populations of positive and negative-signs schizophrenia, or to detect significant behavioral changes on single patients level. The studies disregard behavioral context, namely, what subjects were doing while being recorded, which introduce a lot of noise into collected motor data, and may result in confounds and biases. For example, patients motor behavior may be limited due to external constrains (hospitalization) rather than due to their clinical condition. The measures derived from motor data, including the variability and stability measures take into account the amount of motor activity and the way it changes over time, but do not say anything about the nature of the activity (what type of movement). Therefore, more qualitative measures are needed.

1.2.4 Current Work

In our study we used an accelerometer wearable device, to continuously measure motor behavior of schizophrenia inpatients. The 3-axis accelerometer data was used to derive a rich, quantitative *and* qualitative characterization of patients' behavior over a period of 3 weeks on average. To reduce environmental noise, a substantial part of the analysis focused on times of routine daily activities. Motor data was processed and analyzed in several manners, emphasizing various aspects of patients' behavior, and was then compared with clinical evaluation and drug usage.

In paper 3.1, we describe the general platform of data collection from inpatients in the close ward, and the basic motor features obtained in our study. These include general movement features such as energy and ODBA, as well

as step count and energy variance, which are computed in different activity windows and compared with patients clinical evaluation over time.

In paper 3.2 we describe the process of creating a motor 'codebook', and demonstrate how topic models, a broadly used method for natural language processing (NLP), can be applied to our data in order to achieve a rich, qualitative description of patients' behavior. Using this representation we compare patients behavior to themselves and others, evaluate motor appropriateness and consistency, and characterize clinical sub-populations. In addition, we design a learning algorithm, for automatic patients classification based on motor features and evaluate its performance.

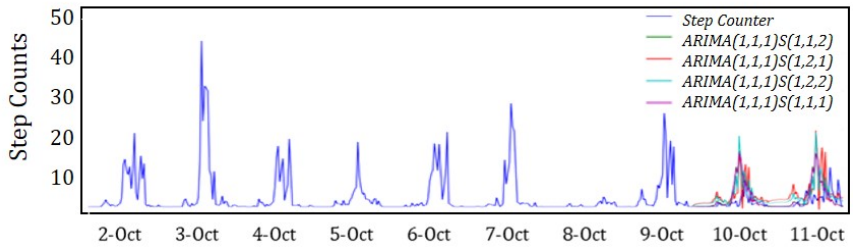


FIGURE 1.7: Illustration of different ARIMA models to predict patient's step counts based on preceding 7 days.

Finally, in paper 3.3, we use a time forecasting method of Auto-Regressive Moving Average (ARIMA) to build a personal model for each patient, which allow us to predict patients motor behavior based on their behavioral patterns in preceding days. We then compare the abnormal behaviors detected by our model with extreme changes in patients' clinical condition as well as specific changes in neuroleptic drug usage (see Figure 1.7).

References

[1] *Affectiva - emotion recognition products*. URL: https://www.affectiva.com/what/products/.

[2] Minu A Aghevli, Jack J Blanchard, and William P Horan. "The expression and experience of emotion in schizophrenia: a study of social interactions". In: *Psychiatry research* 119.3 (2003), pp. 261–270.

[3] Murray Alpert et al. "Prosody and lexical accuracy in flat affect schizophrenia". In: *Psychiatry research* 97.2 (2000), pp. 107–118.

[4] Marco Altini et al. "Combining wearable accelerometer and physiological data for activity and energy expenditure estimation". In: *Proceedings of the 4th Conference on Wireless Health*. ACM. 2013, p. 1.

[5] Keith Anderson and Peter W McOwan. "A real-time automated system for the recognition of human facial expressions". In: *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on* 36.1 (2006), pp. 96–105.

[6] Nancy C Andreasen. "Thought, language, and communication disorders: I. Clinical assessment, definition of terms, and evaluation of their reliability". In: *Archives of general Psychiatry* 36.12 (1979), pp. 1315–1321.

[7] Nancy C Andreasen and Scott Olsen. "Negative v positive schizophrenia: definition and validation". In: *Archives of General Psychiatry* 39.7 (1982), p. 789.

[8] American Psychiatric Association et al. *Diagnostic and statistical manual of mental disorders, (DSM-5®)*. American Psychiatric Pub, 2013.

[9] Hillel Aviezer et al. "The automaticity of emotional face-context integration." In: *Emotion (Washington, D.C.)* 11.6 (Dec. 2011), pp. 1406–14. ISSN: 1931-1516.

[10] Frank J Ayd. "A survey of drug-induced extrapyramidal reactions". In: *Jama* 175.12 (1961), pp. 1054–1060.

[11] Justin T Baker et al. "Toward Expert Systems in Mental Health Assessment: A Computational Approach to the Face and Voice in Dyadic Patient-Doctor Interactions". In: *Iproceedings* 2.1 (2016), e44.

[12] Marian Stewart Bartlett et al. "Recognizing facial expression: machine learning and application to spontaneous behavior". In: *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*. Vol. 2. IEEE. 2005, pp. 568–573.

[13] David R Bassett et al. "Step counting: a review of measurement considerations and health-related applications". In: *Sports Medicine* 47.7 (2017), pp. 1303–1315.

[14] Dror Ben-Zeev et al. "Mobile behavioral sensing for outpatients and inpatients with schizophrenia". In: *Psychiatric services* 67.5 (2015), pp. 558–561.

[15] Jan O Berle et al. "Actigraphic registration of motor activity reveals a more structured behavioural pattern in schizophrenia than in major depression". In: *BMC research notes* 3.1 (2010), p. 149.

[16] Giuseppe Bersani, Elisa Polli, and Giuseppe Valeriani. "Facial expression in patients with bipolar disorder and schizophrenia in response to emotional stimuli: a partially shared cognitive and social deficit of the two". In: *Neuropsychiatric Disease and Treatment* 9 (2013), pp. 1137–1144.

[17] Manfred Bleuler. "Schizophrenia: review of the work of prof. eugen bleuler". In: *Archives of Neurology & Psychiatry* 26.3 (1931), pp. 610–627.

[18] Steve Brown et al. "The unhealthy lifestyle of people with schizophrenia". In: *Psychological medicine* 29.03 (1999), pp. 697–701.

[19] Eric YH Chen et al. "Progressive deterioration of soft neurological signs in chronic schizophrenic patients". In: *Acta Psychiatrica Scandinavica* 102.5 (2000), pp. 342–349.

[20] *Coalition for DSM-5 Reform*. URL: http://dsm5-reform.com/.

[21] Jeffrey F. Cohn et al. "Detecting depression from facial actions and vocal prosody". In: *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops* (Sept. 2009), pp. 1–7.

[22] Paul Ekman. "Universal and cultural differences in facial expression of emotion". In: *Nebraska symposium on motivation*. Vol. 19. University of Nebraska Press Lincoln. 1972, pp. 207–284.

[23] Paul Ekman and Wallace V Friesen. "The repertoire of nonverbal behavior: Categories, origins, usage, and coding". In: *Nonverbal communication, interaction, and gesture* (1981), pp. 57–106.

[24] Eeva A Elliott and Arthur M Jacobs. "Facial expressions, emotions, and sign languages". In: *Frontiers in psychology* 4 (2013).

[25] *Faceware - Facial Motion Capture Software*. URL: http://facewaretech.com/products/software/.

[26] Irina Falkenberg, Mathias Bartels, and Barbara Wild. "Keep smiling!" In: *European archives of psychiatry and clinical neuroscience* 258.4 (2008), pp. 245–253.

[27] Tom FD Farrow et al. "Structural brain correlates of unconstrained motor activity in people with schizophrenia". In: *The British Journal of Psychiatry* 187.5 (2005), pp. 481–482.

[28] Helena Fatouros-Bergman et al. "Stability of facial affective expressions in schizophrenia". In: *Schizophrenia research and treatment* 2012 (2012).

[29] Joseph Firth and John Torous. "Smartphone apps for schizophrenia: a systematic review". In: *JMIR mHealth and uHealth* 3.4 (2015).

[30] Allen Frances. "The new crisis of confidence in psychiatric diagnosis". In: *Annals of internal medicine* 159.3 (2013), pp. 221–222.

[31] Wallace V Friesen and Paul Ekman. "EMFACS-7: Emotional facial action coding system". In: *Unpublished manuscript, University of California at San Francisco* 2 (1983), p. 36.

[32] Wolfgang Gaebel and Wolfgang Wölwer. "Facial expressivity in the course of schizophrenia and depression". In: *European archives of psychiatry and clinical neuroscience* 254.5 (2004), pp. 335–342.

[33] Andrea Gaggioli et al. "A mobile data collection platform for mental health research". In: *Personal and Ubiquitous Computing* 17.2 (2013), pp. 241–251.

[34] Jeffrey M Girard et al. "Social risk and depression: Evidence from manual and automatic facial expression analysis". In: *Automatic Face and Gesture Recognition (FG), 2013 10th IEEE International Conference and Workshops on*. IEEE. 2013, pp. 1–8.

[35] Jeffrey M Girard et al. "Social risk and depression: Evidence from manual and automatic facial expression analysis". In: *Automatic Face and Gesture Recognition (FG), 2013 10th IEEE International Conference and Workshops on*. IEEE. 2013, pp. 1–8.

[36] Paul Gorczynski et al. "Examining strategies to improve accelerometer compliance for individuals living with schizophrenia." In: *Psychiatric rehabilitation journal* 37.4 (2014), p. 333.

[37] Amogh Gudi et al. "Deep learning based facs action unit occurrence and intensity estimation". In: *Automatic Face and Gesture Recognition (FG), 2015 11th IEEE International Conference and Workshops on*. Vol. 6. IEEE. 2015, pp. 1–5.

[38] Hatice Gunes and Maja Pantic. "Automatic, Dimensional and Continuous Emotion Recognition". In: *International Journal of Synthetic Emotions* 1.1 (2010), pp. 68–99.

[39] Daniel Hadar. "Implicit Media Tagging and Affect Prediction from video of spontaneous facial expressions, recorded with depth camera". In: *arXiv preprint arXiv:1701.05248* (2017).

[40] Max Hamilton. "A rating scale for depression". In: *Journal of neurology, neurosurgery, and psychiatry* 23.1 (1960), p. 56.

[41] Jihun Hamm et al. "Automated facial action coding system for dynamic analysis of facial expressions in neuropsychiatric disorders". In: *Journal of neuroscience methods* 200.2 (2011), pp. 237–256.

[42] Jihun Hamm et al. "Dimensional information-theoretic measurement of facial emotion expressions in schizophrenia". In: *Schizophrenia research and treatment* 2014 (2014).

[43] Herman C Hjortsjö. *Man's face and mimic language*. Student litterature, 1969.

[44] *iMotions - Automatic Facial Expression Analysis*. URL: https://imotions.com/facial-expressions/.

[45] Thomas Insel et al. "Research domain criteria (RDoC): toward a new classification framework for research on mental disorders". In: *The American journal of psychiatry* 167.7 (2010), pp. 748–751.

[46] Shashank Jaiswal and Michel Valstar. "Deep learning the dynamic appearance and shape of facial action units". In: *Applications of Computer Vision (WACV), 2016 IEEE Winter Conference on*. IEEE. 2016, pp. 1–8.

[47] Sven Janno et al. "Prevalence of neuroleptic-induced movement disorders in chronic schizophrenia inpatients". In: *American Journal of Psychiatry* 161.1 (2004), pp. 160–163.

[48] Bihan Jiang et al. "A dynamic appearance descriptor approach to facial actions temporal modeling". In: *IEEE transactions on cybernetics* 44.2 (2014), pp. 161–174.

[49] Jyoti Joshi et al. "Can body expressions contribute to automatic depression analysis?" In: *Automatic Face and Gesture Recognition (FG), 2013 10th IEEE International Conference and Workshops on*. IEEE. 2013, pp. 1–7.

[50] Jyoti Joshi et al. "Neural-net classification for spatio-temporal descriptor based depression analysis". In: *Pattern Recognition (ICPR), 2012 21st International Conference on*. IEEE. 2012, pp. 2634–2638.

[51] Samira Ebrahimi Kahou et al. "Emonets: Multimodal deep learning approaches for emotion recognition in video". In: *Journal on Multimodal User Interfaces* 10.2 (2016), pp. 99–111.

[52] Hye Jin Kam et al. "High-Resolution Actigraphic Analysis of ADHD: A Wide Range of Movement Variability Observation in Three School Courses - A Pilot Study." In: *Healthcare informatics research* 17.1 (Mar. 2011), pp. 29–37. ISSN: 2093-369X.

[53] Stanley R Kay, Abraham Flszbein, and Lewis A Opfer. "The positive and negative syndrome scale (PANSS) for schizophrenia." In: *Schizophrenia bulletin* 13.2 (1987), p. 261.

[54] NLW Keijsers et al. "Detection and assessment of the severity of Levodopa-induced dyskinesia in patients with Parkinson's disease by neural networks". In: *Movement disorders* 15.6 (2000), pp. 1104–1111.

[55] Roi Kliper et al. "Evidence for Depression and Schizophrenia in speech Prosody". In: *ISCA*. 2010, pp. 25–27.

[56] Sander Koelstra, Maja Pantic, and Ioannis Patras. "A dynamic texture-based approach to recognition of facial actions and their temporal models". In: *IEEE transactions on pattern analysis and machine intelligence* 32.11 (2010), pp. 1940–1954.

[57] Christian G Kohler et al. "Static posed and evoked facial expressions of emotions in schizophrenia". In: *Schizophrenia Research* 105.1 (2008), pp. 49–60.

[58] Robert LeMoyne et al. "Implementation of an iPhone for characterizing Parkinson's disease tremor through a wireless accelerometer application". In: *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*. IEEE. 2010, pp. 4954–4958.

[59] Stefan Leucht et al. "Second-generation versus first-generation antipsychotic drugs for schizophrenia: a meta-analysis". In: *The Lancet* 373.9657 (2009), pp. 31–41.

[60] Peter F Liddle. "The symptoms of chronic schizophrenia. A re-examination of the positive-negative dichotomy." In: *The British Journal of Psychiatry* 151.2 (1987), pp. 145–151.

[61] Gwen Littlewort et al. "Dynamics of facial expression extracted automatically from video". In: *Image and Vision Computing* 24.6 (2006), pp. 615–625.

[62] Annett Lotzin et al. "Facial emotional expression in schizophrenia adolescents during verbal interaction with a parent." In: *European archives of psychiatry and clinical neuroscience* 263.6 (2013), pp. 529–36.

[63] Patrick Lucey et al. "Automatically detecting pain in video through facial action units". In: *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 41.3 (2011), pp. 664–674.

[64] Simon Lucey, Ahmed Bilal Ashraf, and Jeffrey F Cohn. *Investigating spontaneous facial action recognition through aam representations of the face*. INTECH Open Access Publisher, 2007.

[65] C Mach and S Dollfus. "Scale for assessing negative symptoms in schizophrenia: A systematic review". In: *L'Encephale* 42.2 (2016), pp. 165–171.

[66] Mohammad H Mahoor et al. "A framework for automated measurement of the intensity of non-posed facial action units". In: *Computer Vision and Pattern Recognition Workshops, 2009. CVPR Workshops 2009. IEEE Computer Society Conference on*. IEEE. 2009, pp. 74–80.

[67] Brais Martinez et al. "Automatic analysis of facial actions: A survey". In: *IEEE Transactions on Affective Computing* (2017).

[68] S Mohammad Mavadati and Mohammad H Mahoor. "Temporal Facial Expression Modeling for Automated Action Unit Intensity Measurement". In: *Pattern Recognition (ICPR), 2014 22nd International Conference on*. IEEE. 2014, pp. 4648–4653.

[69] Del D Miller, Stephan Arndt, and Nancy C Andreasen. "Alogia, attentional impairment, and inappropriate affect: their status in the dimensions of schizophrenia". In: *Comprehensive Psychiatry* 34.4 (1993), pp. 221–226.

[70] Arpi Minassian et al. "The quantitative assessment of motor activity in mania and schizophrenia". In: *Journal of affective disorders* 120.1 (2010), pp. 200–206.

[71] Vijay A Mittal et al. "Movement abnormalities and the progression of prodromal symptomatology in adolescents at risk for psychotic disorders." In: *Journal of abnormal psychology* 116.2 (2007), p. 260.

[72] M Morrens et al. "Sensorimotor and cognitive slowing in schizophrenia as measured by the Symbol Digit Substitution Test". In: *Journal of psychiatric research* 40.3 (2006), pp. 200–206.

[73] Manuel Morrens, Wouter Hulstijn, and Bernard Sabbe. "Psychomotor slowing in schizophrenia". In: *Schizophrenia bulletin* 33.4 (2007), pp. 1038–1053.

[74] Jasmine Mote, Barbara K Stuart, and Ann M Kring. "Diminished emotion expressivity but not experience in men and women with schizophrenia." In: *Journal of abnormal psychology* 123.4 (2014), p. 796.

[75] Niamh O'Mahony et al. "Objective diagnosis of ADHD using IMUs." In: *Medical engineering & physics* 36.7 (July 2014), pp. 922–6. ISSN: 1873-4030.

[76] Maja Pantic and MS Bartlett. "Machine analysis of facial expressions". In: *Face Recognition*. Ed. by Kresimir Delac Grgic and Mislav. June. I-Tech Education and Publishing, 2007. Chap. 20, pp. 377–419. ISBN: 9783902613035.

[77] Nicole Papadopoulos et al. "An investigation of gait in children with Attention Deficit Hyperactivity Disorder: a case controlled study". In: *Psychiatry research* 218.3 (2014), pp. 319–323.

[78] S Pappa and P Dazzan. "Spontaneous movement disorders in antipsychotic-naive patients with first-episode psychoses: a systematic review". In: *Psychological medicine* 39.07 (2009), pp. 1065–1076.

[79] R Paradiso et al. "Psyche: personalised monitoring systems for care in mental health". In: *Engineering in Medicine and Biology Society (EMBC), 2010 annual international conference of the IEEE*. IEEE. 2010, pp. 3602–3605.

[80] Victor Peralta and Manuel J Cuesta. "The effect of antipsychotic medication on neuromotor abnormalities in neuroleptic-naive nonaffective psychotic patients: a naturalistic study with haloperidol, risperidone, or olanzapine". In: *Prim Care Companion J Clin Psychiatry* 12.2 (2010), e1–11.

[81] Eduardo Ayala Pérez and Andy Liu Mengxi. "A wearable interface that helps the diagnosis of ADHD behavior based on biofeedback technology". In: *Student Literature* (2012).

[82] P Philippot, R Feldman, and E Coats. *The role of nonverbal behavior in clinical settings.* Ed. by Coats E Philippot P, Feldman R. New York: Oxford University Press, 2003, pp. 3–13.

[83] Harold Alan Pincus, Deborah A Zarin, and Michael First. "Clinical significance and DSM-IV". In: *Archives of general psychiatry* 55.12 (1998), pp. 1145–1145.

[84] Jonathan Posner, James A Russell, and Bradley S Peterson. "The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology". In: *Development and psychopathology* 17.03 (2005), pp. 715–734.

[85] Erik Reinertsen et al. "Continuous assessment of schizophrenia using heart rate and accelerometer data". In: *Physiological Measurement* 38.7 (2017), p. 1456.

[86] James A Russell and Geraldine Pratt. "A description of the affective quality attributed to environments." In: *Journal of personality and social psychology* 38.2 (1980), p. 311.

[87] Georgia Sandbach et al. "Static and dynamic 3D facial expression recognition-A comprehensive survey". In: *Image and Vision Computing* 30.10 (2012), pp. 683–697.

[88] Akane Sano et al. "Recognizing academic performance, sleep quality, stress level, and mental health using personality traits, wearable sensors and mobile phones". In: *Wearable and Implantable Body Sensor Networks (BSN), 2015 IEEE 12th International Conference on*. IEEE. 2015, pp. 1–6.

[89] Arman Savran, BüLent Sankur, and M Taha Bilge. "Comparative evaluation of 3D vs. 2D modality for automatic detection of facial action units". In: *Pattern recognition* 45.2 (2012), pp. 767–782.

[90] Arman Savran, Bülent Sankur, and M Taha Bilge. "Facial action unit detection: 3D versus 2D modality". In: *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*. IEEE. 2010, pp. 71–78.

[91] Arman Savran, Bulent Sankur, and M Taha Bilge. "Regression-based intensity estimation of facial action units". In: *Image and Vision Computing* 30.10 (2012), pp. 774–784.

[92] F Schwab, J Merten, and R Krause. *Expressiveness in dyadic communication*. 1997.

[93] Gwenda Simons et al. "Facial expression in male and female schizophrenia patients." In: *European archives of psychiatry and clinical neuroscience* 260.3 (2010), pp. 267–76.

[94] John Staudenmayer et al. "Methods to estimate aspects of physical activity and sedentary behavior from high-frequency wrist accelerometer measurements". In: *Journal of Applied Physiology* 119.4 (2015), pp. 396–403.

[95] Evelyne Steimer-Krause, Rainer Krause, and Gunter Wagner. "Interaction regulations used by schizophrenic and psychosomatic patients: studies on facial behavior in dyadic interactions". In: *Psychiatry* 53.3 (1990), pp. 209–228.

[96] Ying-li Tian, Takeo Kanade, and Jeffrey F Cohn. "Recognizing action units for facial expression analysis". In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 23.2 (2001), pp. 97–115.

[97] Fabien Trémeau et al. "Facial expressiveness in patients with schizophrenia compared to depressed patients and nonpatient comparison subjects." In: *The American journal of psychiatry* 162.1 (2005), pp. 92–101.

[98] Georgios Tzimiropoulos and Maja Pantic. "Fast algorithms for fitting active appearance models to unconstrained images". In: *International journal of computer vision* 122.1 (2017), pp. 17–33.

[99] Michel Valstar et al. "AVEC 2013 - The Continuous Audio Visual Emotion and Depression Recognition Challenge". In: *cs.nott.ac.uk* (2013).

[100] Supriya Vijay et al. "Computational study of psychosis symptoms and facial expressions". In: *CHI Workshops*. 2016.

[101] Alessandro Vinciarelli, Maja Pantic, and Hervé Bourlard. "Social signal processing: Survey of an emerging domain". In: *Image and Vision Computing* 27.12 (Nov. 2009), pp. 1743–1759. ISSN: 02628856.

[102] John L Waddington. "Schizophrenia, affective psychoses, and other disorders treated with neuroleptic drugs: the enigma of tardive dyskinesia, its neurobiological determinants, and the conflict of paradigms". In: *Int Rev Neurobiol* 31.29 (1989), p. 353.

[103] Sebastian Walther et al. "Objectively measured motor activity in schizophrenia challenges the validity of expert ratings". In: *Psychiatry research* 169.3 (2009), pp. 187–190.

[104] Rui Wang et al. "Studentlife: assessing mental health, academic performance and behavioral trends of college students using smartphones". In: *Proceedings of the 2014 ACM Int Joint Conf on Pervasive and Ubiquitous Computing*. ACM. 2014, pp. 3–14.

[105] Aner Weiss et al. "Toward automated, at-home assessment of mobility among patients with Parkinson disease, using a body-worn accelerometer". In: *Neurorehabilitation and neural repair* 25.9 (2011), pp. 810–818.

[106] Peter F Whitty, Olabisi Owoeye, and John L Waddington. "Neurological signs and involuntary movements in schizophrenia: intrinsic to and informative on systems pathobiology". In: *Schizophrenia bulletin* 35.2 (2009), pp. 415–424.

[107] Adam Wichniak et al. "Actigraphic monitoring of activity and rest in schizophrenic patients treated with olanzapine or risperidone". In: *Journal of psychiatric research* 45.10 (2011), pp. 1381–1386.

[108] Anne-Lise Wolff and Gillian A O'Driscoll. "Motor deficits and schizophrenia: the evidence from neuroleptic-naive patients and populations at risk." In: *Journal of Psychiatry and Neuroscience* 24.4 (1999), p. 304.

[109] Hiroaki Yamamoto et al. "Daily physical activity in patients with schizophrenia". In: *Kitasato Medical Journal* 41 (2011), pp. 145–153.

[110] Amira E Youssef et al. "Auto-optimized multimodal expression recognition framework using 3D kinect data for ASD therapeutic aid". In: *International Journal of Modeling and Optimization* 3.2 (2013), p. 112.

Chapter 2

Facial Analysis Papers

2.1 Automated Facial Expressions Analysis in Schizophrenia: a Continuous Dynamic Approach

Published

Tron, T., Peled, A., Grinsphoon, A., & Weinshall, D. (2015, September). Automated facial expressions analysis in schizophrenia: A continuous dynamic approach. In *International Symposium on Pervasive Computing Paradigms for Mental Health (pp. 72-81).* Springer International Publishing.

Automated Facial Expressions Analysis in Schizophrenia: a Continuous Dynamic Approach

Talia Tron[1,4], Abraham Peled[2,3], Alexander Grinsphoon[2,3], and Daphna Weinshall[4]

[1] Hebrew University, ELSC center for Brain Science, Jerusalem 91904, Israel
[2] Sha'ar Menashe Mental Health Center, Sha'ar Menashe 38706, Israel
[3] Technion Institute of Technology, Rappaport Faculty of Medicine, Haifa 3200003, Israel.
[4] Hebrew University, School of Computer Science and Engineering, Jerusalem 91904, Israel
Talia.Tron@mail.huji.ac.il

Abstract. Facial expressions play a major role in psychiatric diagnosis, monitoring and treatment adjustment. We recorded 34 schizophrenia patients and matched controls during a clinical interview, and extracted the activity level of 23 facial Action Units (AUs), using 3D structured light cameras and dedicated software. By defining dynamic and intensity AUs activation characteristic features, we found evidence for blunted affect and reduced positive emotional expressions in patients. Further, we designed learning algorithms which achieved up to 85% correct schizophrenia classification rate, and significant correlation with negative symptoms severity. Our results emphasize the clinical importance of facial dynamics, and illustrate the possible advantages of employing affective computing tools in clinical settings.

Key words: Schizophrenia, Machine learning, Mental health, Facial expressions, 3D cameras, FACS

1 Introduction

Both clinical observations and computational studies suggest that facial activity plays a major role in signaling people's emotional and mental state [14, 8, 13]. Accordingly, several mental disorders are manifested by reduced or altered facial activity, and facial observations are an integral part of psychiatric diagnosis. To date, there are no objective, quantitative methods to measure these alterations, and no clear relation between them and the underlying brain disturbances. This causes multiple interpretations of phenomenology and results in low reliability and validity of psychiatric diagnosis [2].

Schizophrenia is one of the most severe mental disorders, with lifetime prevalence of about 1% worldwide. The disorder is characterized by negative symptoms, which involve the loss of functions and abilities (e.g. blunted affect), and by positive symptoms, which are pathological functions not present in healthy individuals (e.g. hallucinations). Studies have found that patients with schizophre-

S. Serino et al. (Eds.): MindCare 2015, CCIS 604, pp. 72–81, 20
DOI: 10.1007/978-3-319-32270-4_8

nia demonstrate less positive emotions than controls [10], and lower congruity of emotional response [1]. Furthermore, there has been evidence for reduced upper facial activity [3] and reduced overall facial expressivity [12, 5, 7]. Nonetheless, these studies use a limited set of facial activity characteristic features, not necessarily ecologically relevant, and ignore information regarding facial dynamics and variability. An extensive use of computational methods together with clinical intuition is needed in order to obtain a more comprehensive description of patients behavior.

Our study combines descriptive methods with data-driven analysis. We use machine learning tools and cutting edge technology, in order to study a wide range of facial activity characteristic features, the relation between them, and the way they are manifested in clinical setting.

2 Materials and Methods

2.1 Study Design

Participants The study was done in collaboration with Sha'ar Menashe mental health center. Participants were 34 patients and 33 control subjects. All patients were diagnosed as suffering from schizophrenia according to DSM-5, and the course of illness in these patients varied from 1.5 years up to 37 years, with mean of 16.9 years. All patients but one were under stable drug treatment (mood stabilizer, antidepressant, antipsychotic and/or sedatives). Informed consent was obtained from all individual participants included in the study.

Psychiatric Evaluation Participants were evaluated by a trained psychiatrist using the *Positive and Negative Symptoms Scale* (PANSS), a 30 item scale especially designed to asses the severity of both negative and positive symptoms in schizophrenia [9]. The majority of patients suffered from post-psychotic residual negative signs (Type II) schizophrenia, namely, they showed severe negative symptoms (higher than 5 in the PANSS scale), while severe positive and general symptoms were rather rare (less than 10% of patients). 16 of the symptoms did not vary enough for statistical analysis and learning; therefore, the analysis focused on the remaining symptoms: 3 positive symptoms (Delusions, Conceptual disorganization and Grandiosity), 2 general symptoms (Motor retardation and Poor attention) and 7 negative symptoms (Blunted affect, Emotional withdrawal, Poor rapport, Passive/apathetic social withdrawal, Difficulty in abstract thinking, Lack of spontaneity and flow of conversation and Stereotyped thinking). To test for diagnosis consistency, the PANSS evaluation was repeated independently by a second trained psychiatrist who watched the interview videos. Inter-rater reliability was calculated separately for each PANSS symptom using Pearson correlation test.

Experimental Paradigm All subjects were individually recorded using a 3D structured light camera (carmine 1.09), during a 15 minute long interview con-

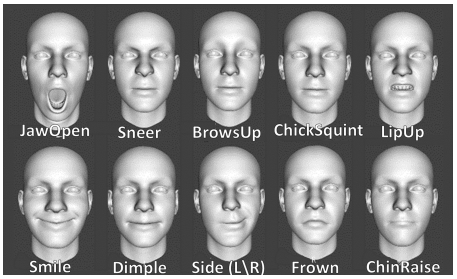
ducted by a trained psychiatrist. The interview was constructed out of one general question ('Tell me about yourself'), and three emotionally evocative questions regarding subject's current mood and recent emotional events. The camera was placed on the table between subject and interviewer, in a way that did not interfere with eye contact and none of the subjects reported discomfort from being recorded. All procedures performed in the study were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

2.2 Facial Activity Features

The Facial Action Coding System (FACS) scores the activity of 46 individual facial muscles called Action Units (AUs) based on their intensity level and temporal segments [4]. Scoring is traditionally done manually, one frame at a time, by certified FACS coders, and automated FACS coding poses a major challenge in the field of affective computing. The advantage of the coding system is that it does not interpret the emotional value of specific features, and allows for a continuous and dynamic facial activity analysis.

Facial Activity Extraction For AUs activity extraction we used the *Faceshift*© commercial software which provides real time 3D face and head tracking, and which is typically used for animating avatars in film and game industry (`www.faceshift.com`). The software automatically analyzes data from 3D cameras based on structured light technology. These cameras capture facial surface data, which is less sensitive to head pose and to lightning conditions than 2D data, and yields a better recognition rate of AUs [11]. *Faceshift* outputs the intensity level over time for 48 AUs. The output was manually evaluated for tracking sensitivity and noise level. Subsequently, 23 *Faceshift* Action Units (AUs) were selected for further analysis and learning, including Brows-up (center, left and right), Mouth-side (left or right), Jaw-open, Lips-up, Lips-Funnel, Eye-In-Right (looking left), Chin-raise, Sneer and both sides (left and right) of Blink, Smile, Frown, Dimple, Lips-Stretch, and Chick-squint (see Fig. 1)

Fig. 1: Illustration of *Faceshift* facial Action Units (AUs) used for learning.



Characteristic Features Computation In order to obtain a detailed characterization of facial behavior, which captures both the dynamics and intensity of the activity in a clinically relevant manner, we calculated 5 characteristic features separately for each AU. First, the raw *Faceshift* signal was quantized using k-means (k=4) clustering. Then a transition matrix was generated, measuring the number of transitions between quantization levels. 5 facial activity characteristic features were then computed:

1. *Activation Ratio* - Fraction of segment during which the AU was activated
2. *Activation Level* - Mean intensity of AU activation
3. *Activation Length* - Number of frames that the AU activation lasted
4. *Change Ratio* - Fraction of the period of AU activation when there was a change in activity level
5. *Fast Change Ratio* - Fraction of fast changes (> 1) in activation level

Activation Level and *Change Ratio* were calculated using frames with non-zero activity only, so that they will not overlap with the *Activation Ratio*. For *Fast Change Ratio*, we normalized the number of fast changes frames by the total number of frames with activity change.

3 Analysis and Learning

The first part of our analysis was descriptive, and was aimed to obtain detailed characterization of facial activity in patients in comparison with controls. In the second part, we applied machine learning tools to generate predictions. We tested whether facial features have predictive power for patients vs. control classification, and for evaluating symptoms severity. To exclude possible confounds such as gender, education level, age and religion, we performed one-way ANOVA; a variable that was found to be different between groups, was further investigated for its effect on facial activity within groups.

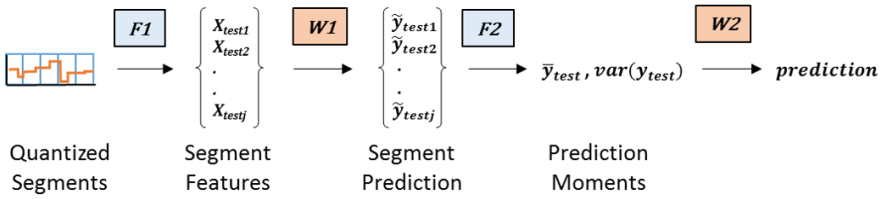
Descriptive Data Analysis In the descriptive part of the analysis, we explored how the facial activity is altered in different parts of the face, paying special attention to smiles. This was done using two tail student's t-tests on the *Activity Level* of each AU separately. For smiles, we further analyzed the difference in all characteristic features, using separate t-test for each feature type. The AU activity was given an emotional interpretation (e.g. high smile level indicates positive emotion), based on the *Emotional Facial Action Coding System* (EMFACS) developed by Paul Ekman, which systematically categorizes combination of AUs to specific emotional categories [6].

To study the way blunted affect is manifested in patients, we performed a regularized ridge regression between symptom severity and all features over all AUs. Feature selection (n=10) was done using f-regression, based on d' scores. Regression results were evaluated by $Pearson's\ R$, and the output regression weights were used for further feature type analysis.

Machine Learning Tools To test the predictive power of our features we trained a learner on train data and evaluated its performance on one test patient at a time, following the Leave-One-Out (LOO) procedure. The basic learning algorithm we used was Support Vector Machine (SVM) for patients vs. control classification, and ridge regression for symptom severity prediction. Before the regression, principle component analysis (PCA) was performed on train data separately for each feature type, resulting in a mixture of AUs. Feature selection was performed based on train data using f-regression (for SVM), or by selecting the highest PCA components (for regression).

To increase learning robustness, we employed a two step prediction algorithm, where each stage is learned separately from train data (see Fig. 2). Interview data of each individual subject was divided into 30 seconds long segments, and 5 representative features were computed separately for each segment ($F1$). In step 1, a learner was trained on the segments of all train subjects, giving as output the first model weights ($W1$) and a prediction for each segment. In step 2, prediction mean and standard deviation over all segments were calculated for each subject ($F2$), and a second learner was trained to predict a participant's label from these moments ($W2$).

Fig. 2: Illustration of the 2-step algorithm.



Performance evaluation was done between-subjects, namely, all segments of one subjects were left out for testing the algorithm. The SVM classifier was evaluated by the area under the Receiver Operator Curve (AUC), a combined measure for the learner's sensitivity (true positive rate) and specificity (true negative rate) with 1 signaling perfect separation and 0.5 signaling chance. Regression results were evaluated by $Pearson's\ R$ between the psychiatrist score and the algorithm prediction, separately for each PANSS symptom.

4 Results

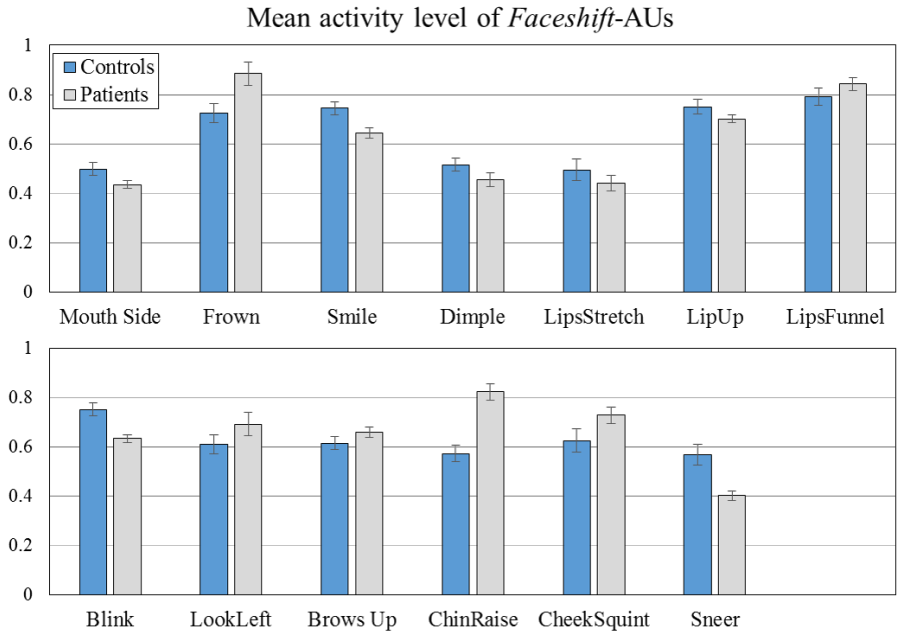
4.1 Inter-rater Reliability

All negative symptoms scores were at high agreement between raters (with an average of $R = 0.850$, $p << 0.01$), and so was 3 positive symptoms ($R = 0.630$, $p = 0.021$ for Delusions, $R = 0.880$, $p << 0.01$ for Conceptual disorganization) and one general symptoms ($R = 0.671$, $p << 0.01$ for Motor Retardation). Poor Attention and Grandiosity were not significantly correlated between raters.

4.2 Facial Activity, Descriptive Analysis

Facial Parts Analysis We found a significant difference in the *Activation Level* of 16 out of 23 *Facefhit*-AUs (see Fig. 3). Specifically, patients demonstrated lower level of activity in Smile, Dimple, Lip-stretch and Lip-up ($p << 0.01$), AUs which are typically in correspondence with positive emotional state. Frowns, Brows-Up and Chin-raise, on the other hand, were at much higher level in patients than in controls, which may indicate the presence of negative valance emotions (sadness, surprise and fear). Although those facial expressions were more intense, they changed more slowly, with reduced *Change Ratio* ($p = 0.004$ for Chin-raise) and *Fast Change Ratio* ($p << 0.01$ for both Chin-raise and Frowns). Blink *Activation Level* was reduced in patients, which in the *Faceshift* framework could mean that they closed their eyes less than controls. Sneer *Activation Level* was also significantly reduced. The level of Cheek-Squint activation was surprisingly enhanced in patients.

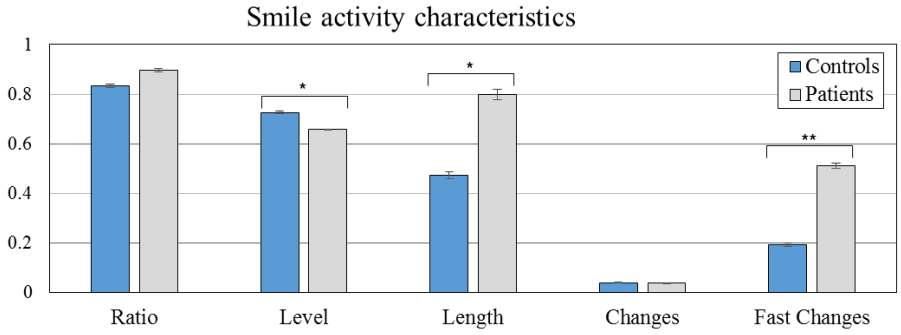
Fig. 3: Mean *Activation level* of facial Action Units in patients and controls. Only significantly different results are presented ($p < 0.05$ in student's t-test).



Smiles Analysis A closer look at smile activation (Fig 4) reveals that in comparison with controls, smile *Activation Level* was reduced, while *Activation Length* and *Fast Change Ratio* were significantly enhanced in patients. These

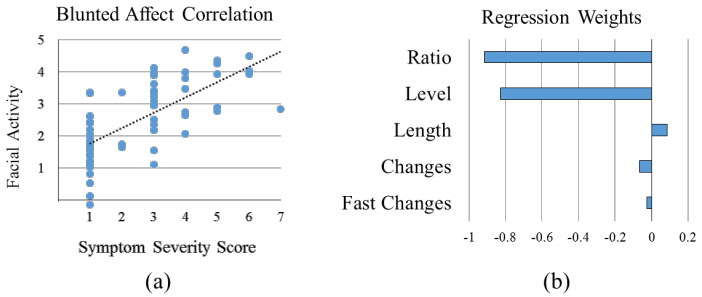
results suggest that in clinical settings, patients may not necessarily smile less, but rather their smiles are at lower intensity, longer, and with faster onset and offset (aka frozen or fake smiles).

Fig. 4: Smile activation characteristic features for patients and controls.



Blunted Affect Regression results (Fig. 5) suggest a significant correlation between AUs activation features and psychiatric evaluation of blunted affect severity ($R_{Pearson} = 0.686$, $p << 0.01$). Based on the regression weights, the two most discriminative AU features were *Activation Level* and *Activation Ratio*, which were in negative correlation with symptom's score. *Change Ratio* and *Fast Change Ratio* were also given negative weights, while *Activation Length* seemed to be positively correlated with the severity of the symptom. These Results are consistent with clinical observations.

Fig. 5: (a) Regression between blunted affect severity and facial activity features. (b) Weights given to each feature by the regression model.

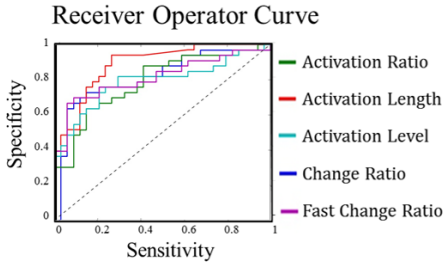


Possible Confounds One-way ANOVA on patients and controls data revealed significant difference between groups for gender ($F = 16.77$, $p << 0.01$) and education level ($F = 6.42$, $p = 0.014$). Neither of these variables was found to have a significant effect on facial activation characteristic within each group. The possible effect of neuroleptic drugs on observed facial activity could not be excluded, since all of our patients were under drug treatment, and additional control is needed.

4.3 Facial Expression Predictive Power

Patients vs. Controls Classification We employed the 2-step learning algorithm one feature type at a time, and using all features together. Each of the feature types was distinctive on its own on test data with AUC significantly better than chance (Fig 6). *Activation Length* gave out the best classification results ($AUC = 0.887$), followed by *Fast change ratio* ($AUC = 0.815$) and *Fast change ratio* ($AUC = 0.814$). This indicates the importance of looking at the the duration and dynamic of facial activity, rather than general intensity measures. The predictive power of using all features together was slightly lower ($AUC = 0.799$), most likely as a result of small sample and subsequent over-fitting.

Fig. 6: (a) ROC curves of each feature type for patients vs. control classification. (b) Classification results summarized as Area Under the ROC Curve (AUC).



(a)

Classification Results

	AUC
Activation Ratio	0.789
Activation Length	0.887
Activation Level	0.790
Change Ratio	0.815
Fast Change Ratio	0.814
All Features	0.799

(b)

PANSS Severity Regression For all negative symptoms, the prediction of the algorithm was significantly correlated with the score given by the psychiatrist ($R > 0.3$, $p \leqslant 0.01$). No such significance was found for any of the positive symptoms, which can be explained by the small variability of positive symptoms scores in our data. We got an unexpected result for general symptoms, with significant correlation only for Poor attention ($R = 0.292$,$p < 0.05$), which outperform the inter-rater correlation for this symptom. Train and test results are summarized in Table 1.

Code	PANSS symptom	Train R	p-val	Test R	p-val
G11	Motor retardation	0.463	1.023E-03	0.154	0.213
G7	Poor attention	0.566	9.35E-07	0.292	0.0166
N1	Blunted affect	0.686	8.27E-10	0.530	4.042E-06
N2	Emotional withdrawal	0.652	4.52E-09	0.510	1.045E-05
N3	Poor rapport	0.550	2.53E-06	0.315	0.00949
N4	Passive/apatheticsocial withdrawal	0.548	2.89E-06	0.368	0.00216
N5	Difficulty in abstract thinking	0.585	3.83E-07	0.369	0.00211
N6	Lack of spontaneity and conversation flow	0.555	1.58E-06	0.301	0.0133
N7	Stereotyped thinking	0.539	3.86E-06	0.369	0.00211
P1	Delusions	0.344	0.005	0.017	0.891
P2	Conceptual disorganization	0.332	0.007	0.065	0.600
P5	Hallucinations	0.306	0.013	0.055	0.660

Table 1: Summary of ridge regression results on train and test data, separately for each PANSS symptom. *Pearson's R* was calculated between the algorithm prediction and symptom severity as scored by a trained psychiatrist.

5 Discussion

Our results are in excellent agreement with previous studies and reported clinical observations. We found clear evidence for clinically reported phenomenon such as blunted affect and lack of positive emotional expressions, and demonstrated how the disorder is manifested differently in different facial parts. Our findings highlight the importance of looking at dynamic characteristics of facial activity and may be employed in clinical settings.

The results give hope that real time automated facial analysis may one day be used for disease monitoring, drug adjustment and treatment outcome evaluation. To achieve these goals, future studies should include monitoring facial activity over time, studying Type-I (positive symptom) schizophrenia patients, and controlling subjects' drug usage. Other future directions include broadening facial activity research to other disorders such as depression and autism, and investigating the relation to neural mechanisms and cognitive performance.

Acknowledgements: This work was supported in part by the Intel Collaborative Research Institute for Computational Intelligence (ICRI-CI), and the Gatsby Charitable Foundations.

References

1. Giuseppe Bersani, Elisa Polli, and Giuseppe Valeriani. Facial expression in patients with bipolar disorder and schizophrenia in response to emotional stimuli: a partially shared cognitive and social deficit of the two. *Neuropsychiatric Disease and Treatment*, 9:1137–1144, 2013.

2. Jeffrey F. Cohn, Tomas Simon Kruez, Iain Matthews, Ying Yang, Minh Hoai Nguyen, Margara Tejera Padilla, Feng Zhou, and Fernando De la Torre. Detecting depression from facial actions and vocal prosody. *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*, pages 1–7, September 2009.
3. Paul Ekman and Wallace V Friesen. The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Nonverbal communication, interaction, and gesture*, pages 57–106, 1981.
4. Paul Ekman and Erika L Rosenberg. *What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS)*. Oxford University Press, 1997.
5. Irina Falkenberg, Mathias Bartels, and Barbara Wild. Keep smiling! *European archives of psychiatry and clinical neuroscience*, 258(4):245–253, 2008.
6. Wallace V Friesen and Paul Ekman. Emfacs-7: Emotional facial action coding system. *Unpublished manuscript, University of California at San Francisco*, 2:36, 1983.
7. Wolfgang Gaebel and Wolfgang Wölwer. Facial expressivity in the course of schizophrenia and depression. *European archives of psychiatry and clinical neuroscience*, 254(5):335–342, 2004.
8. Hatice Gunes and Maja Pantic. Automatic, Dimensional and Continuous Emotion Recognition. *International Journal of Synthetic Emotions*, 1(1):68–99, 2010.
9. Stanley R Kay, Abraham Flszbein, and Lewis A Opfer. The positive and negative syndrome scale (panss) for schizophrenia. *Schizophrenia bulletin*, 13(2):261, 1987.
10. Annett Lotzin, Barbara Haack-Dees, Franz Resch, Georg Romer, and Brigitte Ramsauer. Facial emotional expression in schizophrenia adolescents during verbal interaction with a parent. *European archives of psychiatry and clinical neuroscience*, 263(6):529–36, 2013.
11. Georgia Sandbach, Stefanos Zafeiriou, Maja Pantic, and Lijun Yin. Static and dynamic 3d facial expression recognition: A comprehensive survey. *Image and Vision Computing*, 30(10):683–697, 2012.
12. Gwenda Simons, Johann Heinrich Ellgring, Katja Beck-Dossler, Wolfgang Gaebel, and Wolfgang Wölwer. Facial expression in male and female schizophrenia patients. *European archives of psychiatry and clinical neuroscience*, 260(3):267–76, 2010.
13. Fabien Trémeau, Dolores Malaspina, Fabrice Duval, Humberto Corrêa, Michaela Hager-Budny, Laura Coin-Bariou, Jean-Paul Macher, and Jack M Gorman. Facial expressiveness in patients with schizophrenia compared to depressed patients and nonpatient comparison subjects. *The American journal of psychiatry*, 162(1):92–101, 2005.
14. Michel Valstar, B Schuller, Kirsty Smith, and Florian Eyben. AVEC 2013 - The Continuous Audio/Visual Emotion and Depression Recognition Challenge. *cs.nott.ac.uk*, 2013.

2.2 Facial Expressions and Flat Affect in Schizophrenia, Automatic Analysis from Depth Camera Data

Published

Tron, T., Peled, A., Grinsphoon, A., & Weinshall, D. (2016, February). Facial expressions and flat affect in schizophrenia, automatic analysis from depth camera data. In *2016 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI),* (pp. 220-223). IEEE.

Facial Expressions and Flat Affect in Schizophrenia, Automatic Analysis from Depth Camera Data

Talia Tron[1,4],Abraham Peled[2,3], Alexander Grinsphoon[2,3] and Daphna Weinshall[4]

Abstract— **One of the prominent clinical manifestations of schizophrenia is flat or altered facial activity, and flattening of emotional expressiveness (*Flat Affect*). In this study we used a structured-light depth camera and dedicated software to automatically measure the facial activity of schizophrenia patients and healthy individuals during a short structured interview. Based on K-means clustering analysis, facial activity was characterized in terms of *Typicality*, *Richness* and *Distribution* of 7 facial-clusters. Thus we found patients' facial activity to be poorer, more typical, and characterized mainly by neutral (flat) expressions. The facial features defined in our study achieved up to 85% correct diagnosis classification rate in a SVM based two-step algorithm, and were in significant correlation with *Flat Affect* severity. Our results demonstrate how the use of assistive technology and data-driven computational tools allow for a comprehensive description of patients' facial behavior in clinical settings, and may contribute to the reliability and accuracy of psychiatric diagnosis.**

I. INTRODUCTION

Schizophrenia is one of the most severe mental disorders, with lifetime prevalence of about 1% worldwide. The disorder is characterized by negative symptoms, which involve the loss of functions and abilities (e.g. lack of motivation, cognitive impairments), and by positive symptoms, which are pathological functions not present in healthy individuals (e.g. hallucinations and delusions). Both clinical observations and computational studies suggest that schizophrenia is manifested by reduced or altered facial activity, and by overall affective flattening [15], [14]. *Flat affect*, also known as blunted affect, is clinically defined as 'a severe reduction in emotional expressiveness', and may be expressed in diminished facial expressions, monotonic speech, lack of expressive gestures, and overall apathetic appearance [11]. It is a matter of debate whether the observed flattening is a result of motor or emotional deficits, nonetheless, there is evidence for high congruence between symptom severity, patients wellbeing and treatment outcome [1].

Facial activity is traditionally analyzed in terms of emotional 'prototype expressions' such as anger, fear, sadness, happiness, and disgust [4] in what is known as the *categorical approach* of emotions. Using this approach, it has been shown that patients with schizophrenia demonstrate less positive emotions than controls [12], and lower congruity of emotional response [3]. The downfall of the approach however, is that it uses exaggerated, static and posed facial expressions, while those presented in everyday life are dynamic, spontaneous and far more subtle [9], [2]. An alternative approach is to analyze the facial activity without interpreting its emotional state, which is commonly done using the *Facial Action Coding System* (FACS). This system scores the activity of roughly 46 individual facial muscles called *Action Units* (AUs), based on their intensity level and temporal segments. FACS has been mapped into prototype emotions using the *Emotional Facial Action Coding System* (EMFACS), which systematically categorizes combination of AUs to specific emotions [7] but it can also be used independently. Schizophrenia studies based on FACS has found evidence for reduced upper facial activity [5] and reduced overall facial expressivity [13], [6], [8]. Nonetheless, these studies use a limited set of facial activity characteristic features, not necessarily ecologically relevant, and ignore information regarding facial variability. An extensive use of computational methods together with clinical intuition is needed in order to obtain a more comprehensive description of patients behavior.

Our study suggests a new data-driven approach, combining FACS analysis with the assumption that typical universal emotions can be discovered in a bottom-up analysis. We combine cutting edge technology with data-driven analysis to define a set of 'prototype' facial expression clusters, and to characterize facial activity in terms of *Typicality*, *Richness* and *Distribution* of these clusters. This allow us to study a wide range of facial features, the relation between them, and the way they are manifested in clinical setting.

II. MATERIALS AND METHODS

A. Study Design

The study was done in collaboration with Sha'ar Menashe mental health center. Participants were 34 patients diagnosed as suffering from schizophrenia according to DSM-5 and 33 control subjects. The duration of illness in participating patients was $1.5-37$ years (mean=16.9), and all but one were under stable drug treatment. Informed consent was obtained from all individual participants included in the study.

Participants were individually recorded using a structured-light depth camera (carmine 1.09), during a short structured interview done by a trained psychiatrist which included four questions regarding their emotional state. They then underwent a psychiatric evaluation using the *Positive and*

*This work was supported in part by the Intel Collaborative Research Institute for Computational Intelligence (ICRI-CI), and the Gatsby Charitable Foundations.

[1]ELSC center for Brain Science, Hebrew University, Jerusalem 91904, Israel talia.tron@mail.huji.ac.il

[2]Rappaport Faculty of Medicine, Technion Institute of Technology, Haifa 3200003, Israel

[3]Sha'ar Menashe Mental Health Center, Sha'ar Menashe 38706, Israel

[4]Hebrew University, School of Computer Science and Engineering, Jerusalem 91904, Israel

978-1-5090-2455-1/16/$31.00 ©2016 IEEE

Negative Symptoms Scale (PANSS), a 30 item scale especially designed to assess the severity of both negative and positive symptoms in schizophrenia [10]. All procedures performed in the study were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

B. Facial Activity Extraction

Facial AUs extraction out of depth camera video was done using *Faceshift*© commercial software, which provides real time 3D face and head tracking (`www.faceshift.com`). The software automatically analyzes data from depth-cameras based on structured light technology. The output includes the intensity level over time for 48 facial Action Units (AUs), corresponding to the FACS AUs described in Section I. *Faceshift* output was manually evaluated for tracking sensitivity and noise level, and subsequently 23 AUs were selected for further analysis and learning, including *Brows-up* (center, left and right), *Mouth-side* (left or right), *Jaw-open*, *Lips-up*, *Lips-Funnel*, *Eye-In-Right* (looking left), *Chin-raise*, *Sneer* and both sides (left and right) of *Blink*, *Smile*, *Frown*, *Dimple*, *Lips-Stretch*, and *Chick-squint*.

C. Facial-Cluster Characterization

In order to find the most common combinations of facial-AUs activation in our data, the 23 dimensional vector returned by *Faceshift* was segmented using k-means clustering on data from all subjects simultaneously. Subsequently, each video frame was assigned a cluster label $i \in [k]$ representing its closest cluster centroid (c_i). The resulting facial-cluster centroids can be thought of as the data-driven facial 'prototypes', somewhat equivalent to the categorical expressions described in I, but with no theoretical assumptions regarding the nature of emotions.

The optimal number of clusters was determined using the "elbow criterion". Let V_k be the percent of data variance explained by k centroids. Then $\Delta V = V_k - V_{k-1}$ denotes the difference in the percent of reduced variance when adding one cluster. Under the assumption that ΔV is F distributed, we look for the highest k such that ΔV is statistically significant. In other words, adding more clusters will not significantly improve the ratio of variance explained.

The new vector representation was used to quantitatively describe facial activity in terms of *Richness* (how many prototype expressions appeared), *Typicality* (how similar they were to the prototype) and *Distribution* (which expressions were more prevalent). Facial features were calculated individually for each subject in the following manner:

1) Richness: Let n denote the number of clusters that appeared in a subject's video clip, and k the number of clusters used for the k-means algorithm:

$$Richness = \frac{n-1}{k-1} \tag{1}$$

This measure varies from 0 (only one cluster appeared in the video) to 1 (full richness, all clusters appeared); it can be thought of as a measure for the diversity in facial activity throughout the video.

2) Typicality: Let the Within-Cluster Sum of Squares (*WCSS*) be the sum of distances of each data point x in cluster C_i from its nearest cluster centroid (c_i), with an additional sum over clusters:

$$WCSS_k = \sum_{i=1}^{k} \sum_{\mathbf{x} \in C_i} \|\mathbf{x} - c_i\|^2 \tag{2}$$

For $k = 1$, $WCSS_1$ is proportional to the data variance (the average sqaured distance of the raw data from its mean). For $k > 1$, we define *Typicality* as the percent of the general variance which remains after adding more clusters:

$$Typicality = 1 - \frac{WCSS_k}{WCSS_1} \tag{3}$$

In facial activity terms, we can think of $WCSS_k$ as measuring how similar the video-frame activation is to its assigned 'prototype' among the k facial-clusters. Thus *Cluster Typicality* with score close to 1 indicates that the subject's expressions are similar to the prototypes, while a score close to 0 indicates a significant variability around the prototypes.

3) Cluster Distribution: For each facial-cluster i separately, we counted the number of frames in which it appeared t_i, and normalized it by the length of the video clip T. This allowed for a specific comparison between subjects over the degree of activation of each cluster (or prototype) among the different facial-clusters.

$$Cluster\ Distribution_i = \frac{t_i}{T} \quad \forall i \in [k] \tag{4}$$

D. Data Analysis

Patients vs. controls differences were tested using two-tail student's t-test for *Cluster Typicality*, *Richness*, and *Cluster Distribution* (separately for each facial-cluster). Result significance was evaluated using the *Bonferroni correction*, a family-wise error rate (FWER) for multi-hypothesis testing. In order to allow comparison of our results with the *categorical approach* emotions, the k centroids returned by the clustering algorithm were also evaluated for their affective meaning based on EMFACS (see Section I).

The relation between facial-cluster features and the severity of the *Flat Affect* symptom was tested using regularized ridge regression. A regression model was built for each facial feature separately and for all features together, using a custom designed two-step algorithm (see II-E). *Pearson's* correlation coefficient between the algorithm prediction and the *Flat Affect* score was calculated on train and test data. Symptom's severity was also tested for correlation with all other clinical symptoms scores evaluated by the psychiatrist.

To test for diagnosis consistency, PANSS evaluation was repeated independently by a second trained psychiatrist who watched the interview videos. Inter-rater agreement for *Flat Affect* was tested using *Pearson's R*. Finally, to exclude possible confounds such as gender, education level, age and religion, one-way ANOVA was performed; a variable that was found to be different between groups, was further investigated for its effect on facial activity within groups.

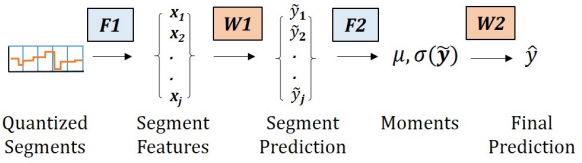


Fig. 1: Illustration of the 2-step algorithm used for learning. Interview data of each individual subject was divided into 30 seconds long segments, and features were calculated separately for each segment ($F1$). In step 1, a learner was trained on the segments of all train subjects, giving as output the first model weights ($W1$) and a prediction for each segment. In step 2 for each subject, prediction mean and standard deviation over all segments were calculated ($F2$) and a second learner was trained to predict subject's label from these moments ($W2$)

E. Learning

In order to evaluate the predictive power of facial-cluster features, we trained a support vector machine (SVM) for patients vs. control classification, and a regularized ridge regression model for *Flat Affect* severity prediction. To increase learning robustness, we employed a two step prediction algorithm, where each stage is learned separately from train data (Fig. 1). The algorithm was trained and tested separately for each feature, and using all features together, following a Leave-One-Out (LOO) procedure with f-regression feature selection (n=5).

Learning performance was evaluated by the Area Under the Receiver Operator Curve (AUC), a combined measure for the learning sensitivity (true positive rate) and specificity (true negative rate) with 1 signaling perfect separation and 0.5 signaling chance. *Pearson's R* was calculated between *Flat Affect* severity score and the algorithm's prediction.

III. RESULTS

A. Facial-Clusters Characteristics

Following the elbow method described in Section II-C, $k = 7$ was chosen for K-means clustering segmentation. Fig. 2 illustrates the centroids of 3 out of 7 facial-clusters returned by the clustering algorithm. The centroid of facial-cluster C_1 (c_1) is characterized by low intensity in all AUs, and may be interpreted as neutral or flat expression. In c_4 we see high intensity of *'ChinLowerRaise'* and *'LipsStretch'*, which correspond with negative valence emotions such as sadness, fear, or anger (according to EMFACS [7]). c_7, on the other hand, is characterized by high intensity smile and dimple, and by overall higher levels of AU activation corresponding to positive emotions such as happiness and content.

TABLE I: Patients vs. controls classification results

Feature Type	*AUC*
Richness and *Cluster Distribution*	0.85
Typicality	0.84
All Features	0.80

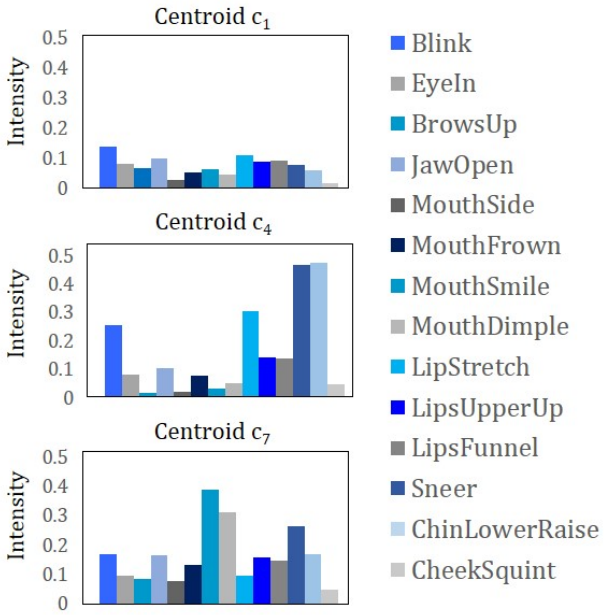


Fig. 2: The centroids of 3 facial-clusters returned by the K-means clustering algorithm (k=7)

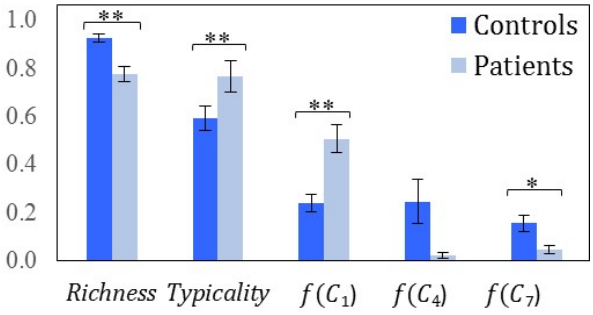


Fig. 3: Group difference for *Richness*, *Typicality*, and for the *Cluster Distribution* (f) of facial-clusters C_1, C_4 and C_7

B. Patients Vs. Control

Significant group differences were found in *Cluster Distribution* for facial-clusters C_1, C_4 and C_7. C_1 was significantly more frequent in patients in comparison with controls ($t = 4.14$, $p << 0.01$), while the frequency of C_4 and C_7 was reduced in patients ($t = 2.43$, $p = 0.018$, and $t = 2.84$, $p = 0.006$ respectively). The results for facial-cluster C_4 are not significant under the *Bonferroni correction*, and further investigation using a larger sample is needed to avoid type-I error. No significant difference was found for the remaining facial-clusters. *Richness* was significantly reduced in patients in comparison with controls ($t = 4.87$, $p << 0.01$), while *Typicality* was higher in patients ($t = -3.39$, $p << 0.01$). Results are summarized in Fig .3.

Learning results suggest that facial-cluster features are predictive for patients vs. control classification (Table I). A classifier (SVM) trained to discriminate between patients and controls, using as input *Richness* and *Cluster Distribution*, achieved the best results ($AUC = 0.85$). *Typicality* achieved

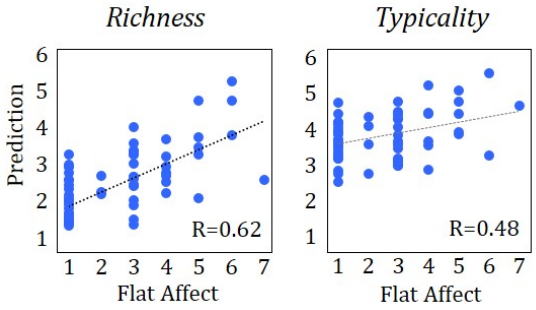


Fig. 4: Correlation between *Flat Affect* score and the prediction of the learning algorithm using the train data

the second best results ($AUC = 0.84$). Classification was not improved by letting the classifier use all the features, most likely due to the small sample limitation and subsequent over-fitting.

C. Correlation with Flat Affect

The evaluation of *Flat Affect* severity was at high agreement between raters ($R = 0.910$, $p << 0.01$), and was found to be significantly correlated with 3 negative symptoms, including *Emotional withdrawal* ($R = 0.907, p << 0$), *Lack of spontaneity and conversation flow* ($R = 0.818, p << 0.01$) and *Difficulty in abstract thinking* ($R = 0.764, p = 0.0014$).

Fig. 4 illustrates the correlation between *Flat Affect* score, and the prediction given by the algorithm based on different feature types. The most highly correlated feature was *Richness*, followed by *Typicality*. Correlation was also significant on test data, ruling out the possibility of mere over-fitting.

Note that the positive correlation is not between symptom severity and feature score, rather it is the correlation with the prediction of the algorithm when learning is *based* on the specific feature. Specifically, the average regression weights ($\bar{w}$) of *Richness* in the first regression ($W1$ in Fig. 1) are negative ($\bar{w} = -0.62$), while *Typicality* is given a positive weight ($w = 0.36$) as expected.

Train and test results are summarize in Table II.

TABLE II: *Pearson* correlation between *Flat Affect* score and algorithm prediction on train and test data

	R-train	p-value	R-test	p-value
All Features	0.647	5.82E-09	0.431	2.72E-04
Richness	0.618	4.18E-08	0.420	3.98E-04
Typicality	0.480	3.912E-05	0.354	0.003
Cluster Distribution	0.472	6.95E-05	0.172	0.163

D. Possible Confounds

One-way ANOVA on patients and controls data revealed significant difference between groups for gender ($F = 16.77$, $p << 0.01$) and education level ($F = 6.42$, $p = 0.014$). Neither of these variables was found to have a significant effect on cluster-facial features. The possible effect of neuroleptic drugs on observed facial activity could not be excluded, since all of our patients were under drug treatment, and additional control is needed.

IV. CONCLUSIONS

Our results are in excellent agreement with clinical findings, and suggest that in clinical settings schizophrenia patients demonstrate a smaller range of expression, characterized mainly by reduced overall facial activity. In contrast to other studies [12], we found a reduction in both positive and negative emotional expressions. Another interesting finding is that *Typicality* is higher in patients. This may indicate that they don't have a different set of basic facial expressions, but rather that their expressivity is less diverse and more repetitive. Finally, we found that information embedded in facial activity is sensitive enough for symptom severity evaluation, and for automatic patient vs. control seperation; this may be one day beneficial for diagnosis, monitoring and treatment.

REFERENCES

[1] Murray Alpert, Stanley D Rosenberg, Enrique R Pouget, and Richard J Shaw. Prosody and lexical accuracy in flat affect schizophrenia. *Psychiatry research*, 97(2):107–118, 2000.

[2] Hillel Aviezer, Shlomo Bentin, Veronica Dudarev, and Ran R Hassin. The automaticity of emotional face-context integration. *Emotion (Washington, D.C.)*, 11(6):1406–14, December 2011.

[3] Giuseppe Bersani, Elisa Polli, and Giuseppe Valeriani. Facial expression in patients with bipolar disorder and schizophrenia in response to emotional stimuli: a partially shared cognitive and social deficit of the two. *Neuropsychiatric Disease and Treatment*, 9:1137–1144, 2013.

[4] Paul Ekman. Universal and cultural differences in facial expression of emotion. In *Nebraska symposium on motivation*, volume 19, pages 207–284. University of Nebraska Press Lincoln, 1972.

[5] Paul Ekman and Wallace V Friesen. The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Nonverbal communication, interaction, and gesture*, pages 57–106, 1981.

[6] Irina Falkenberg, Mathias Bartels, and Barbara Wild. Keep smiling! *European archives of psychiatry and clinical neuroscience*, 258(4):245–253, 2008.

[7] Wallace V Friesen and Paul Ekman. Emfacs-7: Emotional facial action coding system. *Unpublished manuscript, University of California at San Francisco*, 2:36, 1983.

[8] Wolfgang Gaebel and Wolfgang Wölwer. Facial expressivity in the course of schizophrenia and depression. *European archives of psychiatry and clinical neuroscience*, 254(5):335–342, 2004.

[9] Hatice Gunes and Maja Pantic. Automatic, Dimensional and Continuous Emotion Recognition. *International Journal of Synthetic Emotions*, 1(1):68–99, 2010.

[10] Stanley R Kay, Abraham Flszbein, and Lewis A Opfer. The positive and negative syndrome scale (panss) for schizophrenia. *Schizophrenia bulletin*, 13(2):261, 1987.

[11] Peter F Liddle. Schizophrenia: the clinical picture. In *Seminars in general adult psychiatry*, pages 167–186, 2007.

[12] Annett Lotzin, Barbara Haack-Dees, Franz Resch, Georg Romer, and Brigitte Ramsauer. Facial emotional expression in schizophrenia adolescents during verbal interaction with a parent. *European archives of psychiatry and clinical neuroscience*, 263(6):529–36, 2013.

[13] Gwenda Simons, Johann Heinrich Ellgring, Katja Beck-Dossler, Wolfgang Gaebel, and Wolfgang Wölwer. Facial expression in male and female schizophrenia patients. *European archives of psychiatry and clinical neuroscience*, 260(3):267–76, 2010.

[14] Fabien Trémeau, Dolores Malaspina, Fabrice Duval, Humberto Corrêa, Michaela Hager-Budny, Laura Coin-Bariou, Jean-Paul Macher, and Jack M Gorman. Facial expressiveness in patients with schizophrenia compared to depressed patients and nonpatient comparison subjects. *The American journal of psychiatry*, 162(1):92–101, 2005.

[15] Michel Valstar, B Schuller, Kirsty Smith, and Florian Eyben. AVEC 2013 - The Continuous Audio Visual Emotion and Depression Recognition Challenge. *cs.nott.ac.uk*, 2013.

2.3 Differentiating facial incongruity and flatness in schizophrenia, using structured light camera data.

Published

Tron, T., Peled, A., Grinsphoon, A., & Weinshall, D. (2016, August). Differentiating facial incongruity and flatness in schizophrenia, using structured light camera data. In *IEEE 38th Annual International Conference of the Engineering in Medicine and Biology Society (EMBC), 2016* (pp. 2427-2430). IEEE.

Differentiating Facial Incongruity and Flatness in Schizophrenia using Structured Light Camera Data

Talia Tron[1,4], Abraham Peled[2,3], Alexander Grinsphoon[2,3] and Daphna Weinshall[4]

Abstract— **Incongruity between emotional experience and its outwardly expression is one of the prominent symptoms in schizophrenia. Though widely reported and used in clinical evaluation, this symptom is inadequately defined in the literature and may be confused with mere affect flattening. In this study we used structured-light depth camera and dedicated software to automatically measure facial activity of schizophrenia patients and healthy individuals during an emotionally evocative task. We defined novel measures for the congruence of *emotional experience* and *emotional expression* and for *Flat Affect*, compared them between patients and controls, and examined their consistency with clinical evaluation. We found incongruity in schizophrenia to be manifested in a less specific range of facial expressions in response to similar emotional stimuli, while the emotional experience remains intact. Our study also suggests that when taking into consideration affect flatness, no contextually inappropriate facial expressions are evident.**

I. INTRODUCTION

The term schizophrenia was coined by Bleuler in the beginning of the previous century to describe (among other findings) the observed mismatch between patients mood and its outwardly display (*affect*) [3]. In psychiatric clinical evaluation, detecting and monitoring affect incongruity is critically important. First, since this is one of the more typical finding in schizophrenia, it may bias the diagnosis towards it. Second, tracking affect incongruity is a potential candidate for treatment outcome monitoring, since incongruity usually disappears or reduces as the patient positively responds to medications [10].

Currently, the evaluation of affect incongruity is based on the subjective experience of the diagnosing psychiatrist, which leads to low inter-rater reliability and may introduce biases and errors into the diagnosis [13], [6]. Consequently there is a need to define objective measures for mood-congruence assessment, providing psychiatrists with a quantitative indicator for this important symptom.

Although early literature on schizophrenia suggested that the disorder is characterized by the inability to experience pleasure [12], a growing body of work supports the *Inhibition Theory* [3], stating that the *emotional experience* is not compromised, rather that there is an impairment in *emotional expression* [4], [11], [1]. Many studies, together with a substantial amount of clinical observations, report a severe reduction in emotional expressiveness in schizophrenia patients, also known as *Flat* or *Blunted Affect* [2], [6], [9]. Affect flattening may be expressed not only as a reduction in the intensity of emotional expression, but also as a reduction in its variability and dynamics, as demonstrated in our previous study [14].

Flat Affect is typically thought to be independent of affect incongruity and is therefore measured separately using different clinical and empirical scales [15]. Nonetheless, we may wonder whether the clinical and experimental reports of emotional incongruence are not merely a result of misinterpretation of affect flattening. For example, one of Blueler's descriptions of incongruent expression is *'We do not feel his anger even when he speaks of it, because his features and his movements are not in agreement with his words. He may strike us with the most friendly smile on his face'*. Maybe the explanation is that this patient has a smile which doesn't change over time (aka 'frozen smile'), or simply flat affect? A confusion in clinical subjective assessment may arise due to the non-specific definition of affect incongruity.

In the past two decades a great effort has been made to empirically evaluate facial expressions in an objective manner. Most of the studies use the Facial Action Coding System (FACS) which gives scores to the activity of 46 individual facial muscles called Action Units (AUs) based on their intensity level and temporal segments [5]. The advantage of the coding system is that it does not interpret the emotional value of specific features, and allows for a detailed and quantitative facial activity analysis.

In this study we used a structured light camera in order to record schizophrenia patients and healthy individuals while observing emotionally evocative photographs. One benefit of using a depth camera is the resulting ability to accurately track facial motion. This allowed us to obtain fairly reliable measures of AU activity. We further computed discriminative features of the AU activity in order to devise data-driven measures for three phenomenological quantities used in the diagnosis of schizophrenia: *Flat Affect*, *Incongruent Affect*, and *Inappropriate Affect*. We show evidence in our data for the *Flat* and *Incongruent Affects* in schizophrenia paitents, but hardly any *Inappropriate Affect*. We then directly approach the question - can affect incongruity still be observed in schizophrenia patients when taking into account affect

*This work was supported in part by the Intel Collaborative Research Institute for Computational Intelligence (ICRI-CI), and Gatsby Charitable Foundations.

[1]The Edmond and Lily Safra center (ELSC) for Brain Science, Hebrew University, Jerusalem 91904, Israel talia.tron@mail.huji.ac.il

[2]Rappaport Faculty of Medicine, Technion Institute of Technology, Haifa 3200003, Israel

[3]Sha'ar Menashe Mental Health Center, Sha'ar Menashe 38706, Israel

[4]The Rachel and Selim Benin School of Computer Science and Engineering, Hebrew University, Jerusalem 91904, Israel

flatness? our answer in Section III-C is in the affirmative.

II. MATERIALS AND METHODS

A. Study Design

Participants included 34 patients diagnosed with schizophrenia according to the DSM-5 and 33 healthy control subjects matched for age and education level. The course of illness in patients varied from 1.5 years up to 37 years, with mean of 16.9 years. All patients but one were under stable drug treatment (mood stabilizer, antidepressant, anti-psychotic and/or sedatives). Informed consent was obtained from all individual participants included in the study.

Each participant was evaluated by a trained psychiatrist using the *Positive and Negative Symptoms Scale* (PANSS), a 30 item scale especially designed to assess the severity of both negative and positive symptoms in schizophrenia [7]. In addition, they underwent mental status examination, being evaluated for the congruity of their affect (incongruent or congruent).

Subjects were presented with 20 emotionally evocative photos retrieved from the *International Affective Picture System* (IAPS), and were asked to rate their subjective emotional experience while watching the photos ('negative','neutral' or 'positive'). Presenting and rating of the photos followed the IAPS photo rating paradigm [8]. Each subject was individually recorded using a 3D structured light camera (carmine 1.09).

All procedures performed in the study were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

B. Facial Activity Features

Facial activity was extracted out of the video using *Faceshift*©, a commercial software which provides real time 3D face and head tracking (`www.faceshift.com`). The software automatically analyzes data from 3D cameras and outputs the intensity level over time for 48 facial AUs based on the Facial Action Coding System (FACS) described in Section I. This output was manually evaluated for tracking sensitivity and noise level. Subsequently 23 AUs were selected for further analysis and learning, including brows, mouth, jaw, lips, eyes, chick, and nose activation.

For each AU separately, the raw *Faceshift* signal was quantized over time using k-means clustering ($k = 4$), and subsequently 5 different features were calculated. The features were designed to capture clinically relevant characteristics of facial expressions, including the presence of specific AUs, their intensity, and their dynamics (activation length and change ratio). See [14] for a full description of the facial features.

III. AFFECT ANALYSIS AND RESULTS

As stated above, we aim to devise automatic measures of two characteristic manifestations of facial expressions in schizophrenic patients – *Flat Affect* and *Incongruent Affect*. Our analysis below will show that these characteristics are very much correlated in such patients, and therefore the challenge is to device measures which can distinguish the two concepts. In fact, we propose below 3 measures, measuring flatness, incongruity and appropriateness. We show that the 3 measures capture different aspects of the participants' behavior, and specifically they can independently measure *Flat*, *Incongruent* and *Inappropriate* emotional expression.

In Section III-A we analyze a measure of *emotional experience*, which compares the underlying *emotional experience* of participants independently of facial expressions. In Section III-B we propose a measure for the congruence between a person's reported emotional state and facial expressions, which is significantly lower in patients when compared to controls. In Section III-C we directly approach the question whether affect incongruity can still be observed in schizophrenia patients when taking into account affect flatness.

A. Emotional Experience Agreement

Emotional experience is measured independently from emotional expression, by the reported response of participants when asked to describe their subjective emotional state while viewing the photos. Specifically, for each photo separately we calculated the emotional response (negative, neutral or positive) which was most common among all participants in each group (patients and controls). The congruity of emotional experience was measured by the *Agreement Score* - the percentage of participants in the group which reported the most common emotional response. This measure was compared between groups using *Student's t-test*.

Results: The *Agreement Score* was similar in patients and controls (Fig. 1-b), consistent with previous studies which report normal emotional experience in schizophrenia patients (see Section I).

B. Emotional Experience and Facial Expressions

In order to measure the agreement between facial expressions and the emotional experience of participants, we devised the *Predictiveness Score*. For each photo, we extracted the corresponding facial activity segment and computed the corresponding vector of facial activity features (see Section II-B). We excluded photos rated as 'neutral'. For each participant, a linear SVM algorithm was trained to classify photo rating as 'positive' or 'negative' based on the facial features. A Leave One Out (LOO) train-test paradigm was followed, and the *Predictiveness Score* was calculated as the area under the Receiver Operator Curve (AUC), which is a combined measure for classifier specificity and sensitivity. The measure was compared between groups using *Student's t-test* and effect size was evaluated using *Cohen's d'*.

Results: The *Predictiveness Score* was 0.90 for controls, namely, in 90% of the normal population the vector of facial features was useful enough to correctly determine whether the reported emotional experience was positive or negative. This score was significantly lower for patients ($t = 2.829$,

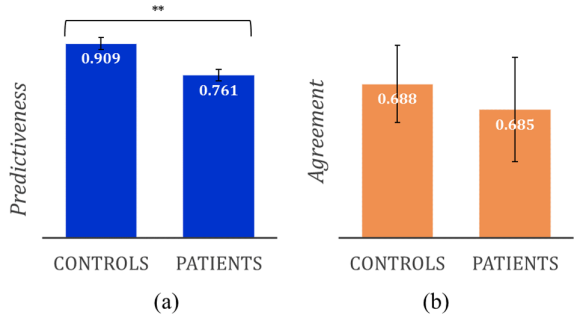


Fig. 1: A difference in emotional expression but not in emotional experience in schizophrenia patients in comparison with controls. a) The *Predictiveness Score* for emotional response classification based on facial expressions was significantly lower in patients ($t = 2.83$, $p < 0.01$). b) No significant group difference was found for the *Agreement Score* between raters.

$p < 0.01$, $df = 30$, effect size=0.508), meaning that in patients facial expressions were much less indicative of their subjective emotional experience (Fig. 1-a).

One should caution against interpreting these results as necessarily demonstrating the incongruity of facial expressions in patients. The difference may be the result of the *Flat Affect* and its effect on the learning procedure, simply because a weaker (lower) signal has typically lower SNR and is therefore harder to learn. In accordance, we found that the clinically *Flat Affect* score is significantly higher in participants with clinically reported incongruity than in congruent participants ($t = 2.08$, $p = 0.045$, $df = 42$ $effectsized' = 0.32$). In order to further isolate the possible influence of affect flatness, we looked at additional measures in the next section.

C. Distinguishing Flatness from Incongruity

In order to define measurements which differentiate facial flatness from facial incongruity, we looked more closely into the signal. We restrict the term *Incongruity* to refer to the consistency of subject's facial expressions when viewing photos evoking a similar emotional response. The term *Inappropriateness* refers to the agreement between subject's facial expressions and the typical expression in the healthy population for photos that elicit a similar emotional response.

Notations: First, the facial feature matrix of all subjects was normalized to Z scores: $Z = [X - E(X)]/\sigma(X)$. Let $j \in [k]$ denote a possible emotional response to a photo, $k = [1 - negative, 2 - neutral, 3 - positive]$. Let n_j denote the set of photos rated as j. Let $N_j = |n_j|$ denote the number of photos rated as j and $N = \sum_j N_j$ denote the total number of rated photos. Let $\mathbf{z}_{ji}$ denote the feature vector of subject's facial activity while watching photo $i \in n_j$. For each participant, $\bar{\mathbf{z}}_j = \frac{1}{N_j}\sum_i \mathbf{z}_{ji}$ denotes the vector of mean facial activity while watching photos in n_j, and $\bar{\mathbf{z}} = \frac{1}{N}\sum_{ji} \mathbf{z}_{ji}$ denotes the vector of mean facial activity while watching any photo.

Flatness Measures: For each subject separately, we calculated the variance in facial activity for similarly rated photos, $SS_{within} = \sum_j \sum_i (\mathbf{z}_{ij} - \bar{\mathbf{z}}_j)^2$, the variance in the average response to differently rated photos $SS_{between} = \sum_j N_j (\bar{\mathbf{z}}_j - \bar{\mathbf{z}})^2$, and the total variance over all photos $SS_{total} = SS_{within} + SS_{between} = \sum_j \sum_i (\mathbf{z}_{ij} - \bar{\mathbf{z}})^2$. Note that all 3 measures may be considered as indicative of general facial flatness, with lower values indicating higher flatness.

TABLE I: Results of *Student's t-test* on our different measures in schizophrenia patients vs. control subjects. SS_{within} and SS_{total} were significantly lower in patients, indicating facial flatness, while *Incongruity* tended to be higher in patients. *Inappropriateness* was not significantly different in the two populations.

	Controls	Patients	t	p-val
SS_{within}	7.52	15.91	2.305	0.027
SS_{total}	8.53	18.07	2.376	0.022
Incongruity	0.92	0.88	-2.014	0.051
Inappropriateness	28.62	38.27	1.414	0.168

Results: The total facial activity variance SS_{total} was higher in controls in comparison with patients ($t = 2.38$, $p = 0.022$, $effectsized' = 0.43$), and so were the other types of variance ($t = 2.30$, $p = 0.027$, $effectsized' = 0.41$ for SS_{within} and $t = 2.16$, $p = 0.037$, $effectsized' = 0.39$ for $SS_{between}$), see Table I. This reinforces the clinical observation that patients' affect is more flat than that of controls.

Congruity Measures: *Incongruity* is defined as the ratio between the variance within each emotion and the total variance (1). This measure tells us how unique facial expressions are with respect to each emotional state relative to the overall variance in facial expressions. In other words, we measure the inconsistency of emotional response given the total emotional flatness.

$$Incongruity = \frac{SS_{within}}{SS_{total}} \tag{1}$$

Inappropriateness is defined as the difference between the typical expression in response to a specific stimuli (the mean of facial activity over all control subjects), and each subject's individual facial activity (2). This tells us how different subject's expressions are from the expected normal response.

Specifically, let $\bar{\mathbf{z}}_{Hj}$ denote the mean facial activity of control subjects while watching photo j. For each subject, $\bar{\mathbf{z}}_{Hj}$ was calculated using all remaining control subjects (following a LOO paradigm). Then:

$$Inappropriateness = \sum_j \sum_i (\mathbf{z}_{ij} - \bar{\mathbf{z}}_{Hj})^2 \tag{2}$$

We also define a normalized version of this same measure, dividing each summand of least square distances by the variance of the subject's response to similar photos: $\sum_j [\sum_i (\mathbf{z}_{ij} - \bar{\mathbf{z}}_{Hj})^2 / \sum_i (\mathbf{z}_{ij} - \bar{\mathbf{z}}_j)^2]$. In each of the measures listed above, we tested the difference between schizophrenia patients and control subjects using *Student's t-test*.

Results: We found that *Incongruity* tend to be higher in patients ($t = -2.014$, $p = 0.051$, $df = 30$, $effectsized' = 0.36$), indicating that the facial response to stimuli evoking similar emotions is less consistent. The *Inappropriateness* measure did not significantly differ between groups nor did the normalized measure. Results are summarized in Table I.

Correlation with Clinical Evaluation: To justify the interpretation of our measures in terms of affect congruity and facial flatness we further tested the way these measures behave in accordance with clinical evaluation. Correlation between the various measures and the clinical score for

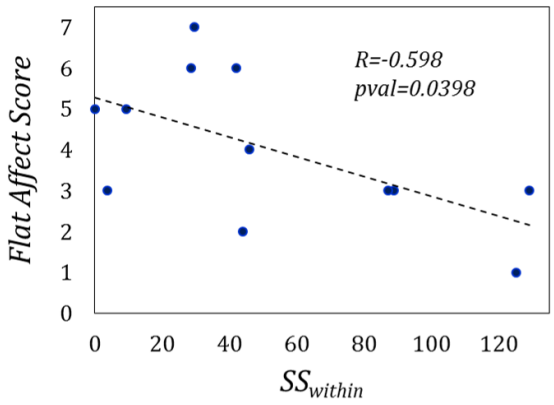


Fig. 2: *Pearson* correlation between patients' variance in facial expressions while looking at photos which elicit similar emotional response (SS_{within}), and the *Flat Affect* score given to them by the psychiatrist (based on the PANSS scoring system).

Flat Affect was calculated, in the patients group only, using *Pearson correlation test*. The difference in these measures between subjects diagnosed as having incongruent affect vs. congruent subjects was evaluated using *Student's t-test*.

Results: (i) *SSwithin* was found to be negatively correlated with the *Flat Affect* score given by the psychiatrist ($R = -0.598$, $p = 0.0398$, Fig. 2), as expected, supporting our interpretation of this measure as indicative of the clinical *Flat Affect*. No such correlation was found for the *Incongruity* measure ($r = -0.0066$, $p = 0.980$). (ii) The *Incongruity* measure was found to be significantly higher in subjects that were clinically diagnosed as having incongruent affect ($t = 2.238$, $p = 0.0373$, $df = 22$, $effectsized' = 0.47$). No significant difference was found for any of the other measures. Additionally, the *Incongruity* measure was not correlated with the *SSwithin* measure ($r = 0.10$, $p = 0.585$). These results reinforce our choice of measures and are in accordance with their clinical interpretation.

IV. SUMMARY AND DISCUSSION

We automatically tracked facial activity of schizophrenia patients and healthy individuals during an experimental paradigm where subjects watched a set of emotionally evocative photos and reported their emotional response. From the recorded facial activity, we obtained automatic measures of two phenomenological characteristics of schizophrenia: *Flat Affect*, which describes a severe reduction in emotional expressiveness, and *Incongruent Affect*, which refers to a mismatch between facial expression and the subjective emotional response to the current situation. Our study refined and readjusted the definition of the two symptoms, eliminating the prevalent misinterpretation of facial flatness as incongruity. Our results suggest that *Incongruity* is manifested independently of affect flatness, by means of a less specific range of responses to similarly evocative emotional stimuli.

Specifically, we showed that affect flatness is measured reliably by the variance in facial expressions, while the relative variance in response to similar emotional stimuli is a separate, reliable measure of affect incongruity. When taking into consideration affect flatness, we did not find evidence for *Inappropriate Affect*, when the emotional response does not fit the socially expected response in similar situations. The measures introduced in our study, together with their clinical interpretations, should be further examined in both clinical and empirical settings in light of these results.

Theoretically, our results support the *inhibition hypothesis* of Kraepelin and Bleuler [3], which suggests that impairment and blunting of facial expression as observed in schizophrenia patients is not a result of compromised emotional experience. We found the emotional experience of schizophrenia patients to be comparable to that of healthy individuals.

The study demonstrates the importance of an objective, quantitative and precise definition of clinical symptoms in schizophrenia. Our hope is that the novel measures and analysis approach we offer will contribute to the reliability of psychiatric diagnosis, will allow better characterization of patients' behavior, and will promote both research and treatment in the field.

REFERENCES

[1] Minu A Aghevli, Jack J Blanchard, and William P Horan. The expression and experience of emotion in schizophrenia: a study of social interactions. *Psychiatry research*, 119(3):261–270, 2003.
[2] Giuseppe Bersani, Elisa Polli, and Giuseppe Valeriani. Facial expression in patients with bipolar disorder and schizophrenia in response to emotional stimuli: a partially shared cognitive and social deficit of the two. *Neuropsychiatric Disease and Treatment*, 9:1137–1144, 2013.
[3] Manfred Bleuler. Schizophrenia: review of the work of prof. eugen bleuler. *Archives of Neurology and Psychiatry*, 26(3):610, 1931.
[4] Alex S Cohen and Kyle S Minor. Emotional experience in patients with schizophrenia revisited: meta-analysis of laboratory studies. *Schizophrenia bulletin*, 36(1):143–150, 2010.
[5] Paul Ekman and Erika L Rosenberg. *What the face reveals- Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS)*. Oxford University Press, 1997.
[6] Jihun Hamm, Amy Pinkham, Ruben C Gur, Ragini Verma, and Christian G Kohler. Dimensional information-theoretic measurement of facial emotion expressions in schizophrenia. *Schizophrenia research and treatment*, 2014, 2014.
[7] Stanley R Kay, Abraham Flszbein, and Lewis A Opfer. The positive and negative syndrome scale (panss) for schizophrenia. *Schizophrenia bulletin*, 13(2):261, 1987.
[8] Peter J Lang, Margaret M Bradley, and Bruce N Cuthbert. International affective picture system (iaps): Affective ratings of pictures and instruction manual. *Technical report A-8*, 2008.
[9] Christian Lindner, Udo Dannlowski, Jochen Bauer, Patricia Ohrmann, Rebekka Lencer, Pienie Zwitserlood, Harald Kugel, and Thomas Suslow. Affective flattening in patients with schizophrenia: Differential association with amygdala response to threat-related facial expression under automatic and controlled processing conditions. *Psychiatry investigation*, 13(1):102–111, 2016.
[10] Herbert Y Meltzer, Alison A Sommers, and Daniel J Lunchis. The effect of neuroleptics and other psychotropic drugs on negative symptoms in schizophrenia. *Journal of clinical psychopharmacology*, 6(6):329–338, 1986.
[11] Jasmine Mote, Barbara K Stuart, and Ann M Kring. Diminished emotion expressivity but not experience in men and women with schizophrenia. *Journal of abnormal psychology*, 123(4):796, 2014.
[12] Sandor Rado. Dynamics and classification of disordered behavior. *American Journal of Psychiatry*, 110(6):406–416, 1953.
[13] Hanno Santelmann, Jeremy Franklin, Jana Bußhoff, and Christopher Baethge. Test–retest reliability of schizoaffective disorder compared with schizophrenia, bipolar disorder, and unipolar depressiona systematic review and meta-analysis. *Bipolar disorders*, 17(7):753–768, 2015.
[14] Talia Tron, Abraham Peled, Alexander Grinsphoon, and Daphna Weinshall. Automated facial expressions analysis in schizophrenia: a continuous dynamic approach.
[15] Paula T Trzepacz and Robert W Baker. *The psychiatric mental status examination*. Oxford University Press, 1993.

Chapter 3

Motor Analysis Papers

3.1 Real-time Schizophrenia Monitoring using Wearable Motion Sensitive Devices

Published

Tron, T., Resheff, Y. S., Bazhmin, M., Peled, A., & Weinshall, D. (2017, November). Real-time Schizophrenia Monitoring using Wearable Motion Sensitive Devices. In *the 7th EAI International Conference on Wireless Mobile Communication and Healthcare (MobiHealth)*.Springer, Berlin, Heidelberg.

Real-time Schizophrenia Monitoring using Wearable Motion Sensitive Devices*

Talia Tron[1], Yehezkel S. Resheff[1], Mikhail Bazhmin[2], Abraham Peled[2,3], and Daphna Weinshall[4]

[1] The Edmond and Lily Safra center (ELSC) for Brain Science, Hebrew University of Jerusalem (HUJI), Israel
[2] Sha'ar Menashe Mental Health Center, Hadera Israel
[3] Rappaport Faculty of Medicine, Technion Institute of Technology, Haifa Israel
[4] The Rachel and Selim Benin School of Computer Science and Engineering, HUJI

Summary. Motor peculiarity is an integral part of the schizophrenia disorder, having various manifestations both throughout the phases of the disease, and as a response to treatment. The current subjective non-quantitative evaluation of these traits leads to multiple interpretations of phenomenology, which impairs the reliability and validity of psychiatric diagnosis. Our long-term objective is to quantitatively measure motor behavior in schizophrenia patients, and develop automatic tools and methods for patient monitoring and treatment adjustment. In the present study, wearable devices were distributed among 25 inpatients in the closed wards of a Mental Health Center. Motor activity was measured using embedded accelerometers, as well as light and temperature sensors. The devices were worn continuously by participants throughout the duration of the experiment, approximately one month. During this period participants were also clinically evaluated twice weekly, including patients' mental, motor, and neurological symptom severity. Medication regimes and outstanding events were also recorded by hospital staff. Below we discuss the general framework for monitoring psychiatric patients with wearable devices. We then present results showing significant correlations between features of activity in various daily time-windows, and measures derived from the psychiatrist's clinical assessment or abnormal events in the patients' routine.

1 Introduction

The relevant clinical literature describes a wide range of motor pattern alternations, manifested in different phases of the schizophrenia disorder. Positive-signs schizophrenia patients are typically psychotic and disorganized, characterized mainly by positive symptoms (e.g. auditory hallucinations, delusions and paranoid thoughts). In clinical settings, these patients show invol-

* Published in Proc. of 7th EAI International Conference on Wireless Mobile Communication and Healthcare (MobiHealth), Nov 2017, Vienna Austria.

untary movements, dyskinesia and catatonic symptoms [1]. In negative-signs schizophrenia, there is usually an observed motor retardation, psycho-motor poverty, decreased spontaneous movements, psycho-motor slowing and flattened affect [2, 3]. Some patients demonstrate both types simultaneously or during different phases of the illness.

Neurological Soft Symptoms (NSS) can manifest early and during the progression of the disorder, and include deficits in coordination, sensory integration, and sequential motor behaviors [4]. Medical treatment was found to improve some of the motor symptoms, including NSS, involuntary movement and dyskinesia [1]. These medications, however, may also introduce in chronic patients drug-induced movement disorders such as tremor dystonia, Parkinsonism (rigidity and bradykinesia), akathisia and tardive dyskinesia [5].

The diversity and specificity of motor symptoms throughout different phases of the disorder and as a response to drugs, makes them good candidates for patient monitoring and treatment outcome evaluation. Nonetheless, to date, these symptoms are evaluated in a descriptive non etiological manner based on subjective clinical scales such as the Unified Dyskinesia Rating Scale (UDysRS) [6] and the Unified Parkinson's Disease Rating Scale (MDS-UPDRS) [7]. The lack of objective, quantitative methods to measure these symptoms, and the insufficient conceptual clarity around it, causes multiple interpretations of phenomenology, often entailing low reliability and validity of the diagnosis. In addition, symptom evaluation process requires expert staff and availability of resources, and it is not done frequently enough to capture delicate changes in patients' spontaneous and drug-induced conditions.

The last decade has seen a steep rise in the use of wearable devices in medical fields ranging from human physiology [8] to movement disorders [9,10] and mental health [11]. Accelerometers and gyroscopes, which are commonly embedded in smart-watches and other wearable devices, are now used to assess mobility, recognize activity, and context. In a clinical setting, these sensors may be used in order to detect change in high-level movement parameters, track their dynamics and correlate them with mental state.

The objective of the current study is to develop and evaluate a framework, where wearable devices are used to facilitate continuous motor deficits monitoring in schizophrenia patients in a natural setting. This is an important step towards a detailed automatic evaluation system of symptom severity in schizophrenia. Such a system has a great potential to help understand this illusive disease. An additional goal would be to help with the overwhelming need for detection and characterization of sub-types of the disease towards a better understanding of underlying causes, and the development of better and more personalized treatment.

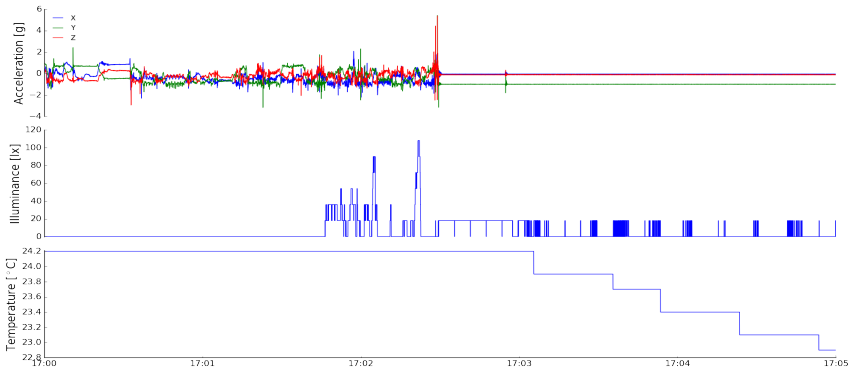


Fig. 1. Raw data as recorded by the smart-watches, including tri-axial accelerometer (top panel), light sensor (middle), and temperature (bottom). This plot shows data from a single patient, recorded on 28 Jan, 2017 at 5:00-5:05pm. Only accelerometer data was used for further analysis.

2 Methods

2.1 Participants and clinical evaluation

Twenty seven inpatients from the closed wards at Shaar-Meashe MHC participated in the study after signing appropriate Helsinki legal consents. Most participants (21/27) were diagnosed with schizophrenia according to the DSM-5, 3 with paranoid schizophrenia, 2 with schizoaffective disorder, and one with psychotic state cannabinoids. Participants' age varied from 21 to 58 (mean of 37.48), with course of illness varying from 0 (first hospitalization) up to 37 years (mean of 16.9 years). Two of the patients dropped out of the study after less than a day due to lack of cooperation. The rest (25 patients) were followed for a period of three weeks on average (6-52 days).

The study was conducted in natural settings, where patients were *not* required to change any personal or medical procedure. In addition to routine reports by nurses and physicians, every patient underwent an additional evaluation by a trained psychiatrist twice a week. The procedure included medication monitoring (type, dosage and frequency), as well as clinical evaluation of positive and negative symptom severity (PANSS [12]) and neurological symptoms severity (NES [13]).

All procedures performed in the study were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

2.2 Data Acquisition

At study onset, each participant was given a smart-watch (GeneActiv[5]). The devices included tri-axial accelerometers, light, and temperature sensors, the high frequency output ($50Hz$) of which was stored on memory cards embedded in the device (see Fig. 1). Data was collected by the aforementioned smart watches worn continuously by patients throughout the experiment (for a total of 489 days of data from 25 patients). The devices were placed and removed by the medical staff, and the content of the memory card was uploaded to a central storage location upon termination of the experiment for further analysis.

In order to reduce noise introduced by the variability in patients activity which is due to external circumstances rather than mental state, weekends were excluded from the study and our analysis focused on fixed time windows with regular departmental daily activity. These included occupational therapy time slots (10am-11am), lunch (12pm-1pm), and indoor free time (4pm-5pm). In addition, we calculated full day features (6am-10pm) and used night time features (10pm-6am) to evaluate sleep quality.

2.3 Features

Features were computed on the basis of the accelerometer readings, analyzed in 1 minute windows (see Table 1 and Fig. 2). Light and temperature data were not used for the analysis. The point-wise sum of values and sum of square values of the tri-axial accelerometer measurements (Energy Square and Energy Sum respectively) were averaged over 1 minute intervals. The variance of the sum of squares (Energy Variance) was also computed over the same window. Stepping behavior (Step Detector) was detected as large maxima of the smoothed square norm of the point-wise acceleration. Overall Dynamic Body Acceleration (ODBA), a measure of energy expenditure, was computed as the mean norm of the accelerometer signal after application of a high-pass filter.

Table 1. List of features calculated on the basis of the tri-axial Accelerometers. Average and variance was calculated on a 1 minute time window.

Feature	Description
Step Detector	Simple count of the number of steps per minute
Energy Square	Averaged sum of point-wise square acceleration
Energy Sum	Averaged sum of point-wise acceleration
Energy Variance	Variance of point-wise square acceleration
ODBA	Mean norm of a high-passed version of acceleration

[5] https://www.geneactiv.org/

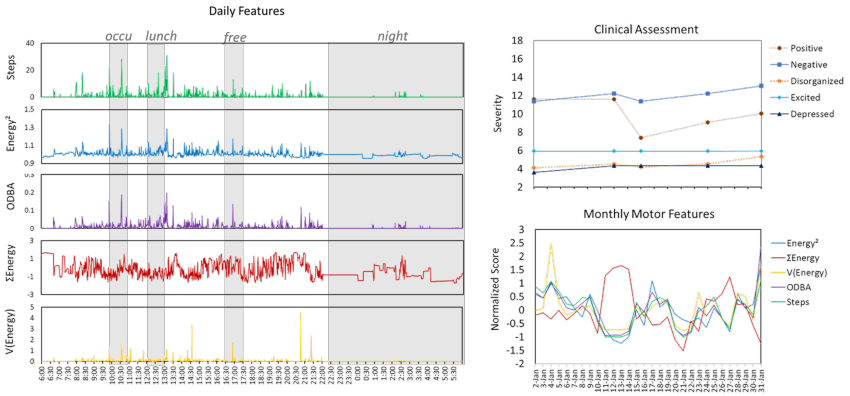


Fig. 2. The daily features of a single subject (left): gray areas indicate the time windows used for aggregated feature calculation. Monthly follow-up of a single patient (right): top panel shows the clinical five-factor PANSS score given by a trained psychiatrist on a bi-weekly basis; bottom panel shows the aggregated features calculated based on the different time windows.

2.4 Clinical Assessments

The 30-item scale for positive and negative symptom assessment (PANSS) was reduced to the following 5 literature-based factors: Positive, Negative, Disorganized/Concrete, Excited and Depressed. The dimensionality reduction was done according to the consensus model suggested by Wallwork et al. [14], based on 25 previously published models and refined with confirmatory factor analysis (CFA).

The negative and positive factors had low between-factor correlation ($R = 0.399$), indicating good separation of the symptomatology space. As expected, the positive factor was in high correlation with the mean of all positive PANSS items ($R = .944$), and likewise the negative factor was in high correlation with the mean of all negative PANSS items ($R = .972$).

3 Results

We investigated two distinct ways by which wearable devices can be used for patient monitoring, in order to assist physicians in understanding the state of a patient. The first aspect of monitoring relates to the automatic assessment of a patient's condition, in order to provide automated, continuous, and objective measures of mental state. To this end we investigated the correlation between the computed measures and assessments by physicians, as described in Section 3.1. The second aspect of monitoring relates to the detection of change (or anomalous behavior patterns) which warrants additional attention from the medical staff, as described in Section 3.2.

3.1 Movement patterns and mental state

In order to investigate the correspondence between patterns of movement and mental state, *multiple correlation analysis* was computed between activity related features (described in Section 2.2) and PANSS factors. Results (Table 2) indicate the predictive benefit of the computed activity-related features with respect to the PANSS factors. When separately considering features computed in each of the time-windows, it is evident that different time windows provide varying predictive value for the 5 different PANSS factors.

Table 2. Percent Explained Variance based on *Multiple Correlation* between computed features in each of the 5 time-windows and each of the 5 PANSS factors. (See Section 2.2 for time-window specifications.)

	free	lunch	occu	day	night	all
Positive Factor	16.30%	11.14%	12.31%	19.80%	5.21%	53.77%
Negative Factor	19.74%	3.15%	2.06%	18.36%	9.77%	55.50%
Disorganized/Concrete Factor	22.73%	0.50%	15.13%	13.42%	5.82%	64.81%
Excited Factor	23.79%	8.75%	15.08%	10.35%	12.70%	57.10%
Depressed Factor	31.01%	9.23%	8.94%	5.78%	6.39%	58.33%

Specifically, the Depressed Factor is described relatively well using features from the *free time* window, with 31.01% explained variance, while all other time-windows are below 10%. Both Positive and Negative factors are described well using features from the *free time* as well as *all day* time-windows. The remaining factors are again best described using *free time*. Overall, the *free time* window is the single most effective window, presumably since it imposes less structure on the movement of the subjects, allowing for the manifestation of the underlying mental state. In all cases, combining all time windows (rightmost column in Table 2) leads to substantially higher explained variance, compared to any of the individual windows.

Interestingly, looking at individual variable correlations we see that step count during *free time* was positively correlated with positive, disorganized and exited factor ($R = 0.37, 0.37$ and 0.31 respectively), but not with the negative and depressed factors. In addition, patients who had higher scores in disorganized and exited factors tended to have lower Energy scores during occupational time ($R = -0.28$ for Energy Sum and -0.22 for Energy Variance). This may indicate some motor retardation which is manifested only in non-walking time.

3.2 Continuous Monitoring

Our measures can be used to track changes in the patient's condition as compared to some established normal baseline, and may identify external events

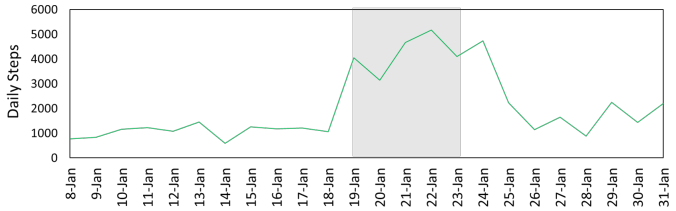


Fig. 3. Mean *daily steps* of a single subject. The gray area corresponds to a short-lasting change in medication regime.

which are correlated with the departure from normality. Fig. 3 demonstrates such a case: daily *step counts* of a patient dramatically increased 5-fold, at the same time as a significant change in medication dosage was introduced. Whether the change in medication *caused* the rise in movement propensity or they were both triggered by a change in mental state, this observation points to the relevance of monitoring macro movement patterns as part of routine patient monitoring.

4 Conclusions

We describe a study designed to evaluate the utility of wearable devices fitted with accelerometer, light, and temperature sensors, for the monitoring of schizophrenia patients in a closed ward mental health institution. Initial results show correlations between features of activity in various daily time-windows, and factors derived from the PANSS assessment.

Results indicate that movement features during free time are the most indicative of mental state. This finding is somewhat counter-intuitive, since the more structured activity during occupational therapy or lunch was expected to highlight differences in the state of patients. However, our results clearly show that the behavior of individuals when left to their own devices is better correlated with the PANSS factors.

These findings points to the possibility of automatically and continuously tracking Schizophrenia related symptoms and patient state, in a natural setting hospital environment.The benefits of such a tracking system are twofold; first, the continuous tracking will assist physicians in understanding the state of a patient on an on-going basis, as opposed to specific points in time, when assessed by the doctor. Second, long term monitoring of a large number of patients will produce data which will allow us to develop more objective measures of the motor aspects of the illness, and facilitate a more personalized, objective, and data driven approach which is much needed in the field of mental health.

Future work will focus on measuring the utility of this approach as an augmentation tool from a physicians perspective on the one hand, and the

ability to predict physician assessments for automation of diagnosis on the other.

References

1. Victor Peralta and Manuel J Cuesta. The effect of antipsychotic medication on neuromotor abnormalities in neuroleptic-naive nonaffective psychotic patients: a naturalistic study with haloperidol, risperidone, or olanzapine. *Prim Care Companion J Clin Psychiatry*, 12(2):e1–11, 2010.
2. Manuel Morrens, Wouter Hulstijn, and Bernard Sabbe. Psychomotor slowing in schizophrenia. *Schizophrenia bulletin*, 33(4):1038–1053, 2007.
3. Adam Wichniak, Agnieszka Skowerska, Jolanta Chojnacka-Wójtowicz, Tomasz Taflíński, Aleksandra Wierzbicka, Wojciech Jernajczyk, and Marek Jarema. Actigraphic monitoring of activity and rest in schizophrenic patients treated with olanzapine or risperidone. *Journal of psychiatric research*, 45(10):1381–1386, 2011.
4. Igor Bombin, Celso Arango, and Robert W Buchanan. Significance and meaning of neurological signs in schizophrenia: two decades later. *Schizophrenia Bulletin*, 31(4):962–977, 2005.
5. Sven Janno, Matti Holi, Katinka Tuisku, and Kristian Wahlbeck. Prevalence of neuroleptic-induced movement disorders in chronic schizophrenia inpatients. *American Journal of Psychiatry*, 161(1):160–163, 2004.
6. Christopher G Goetz, John G Nutt, and Glenn T Stebbins. The unified dyskinesia rating scale: presentation and clinimetric profile. *Movement Disorders*, 23(16):2398–2403, 2008.
7. Christopher G Goetz, Barbara C Tilley, Stephanie R Shaftman, Glenn T Stebbins, Stanley Fahn, Pablo Martinez-Martin, Werner Poewe, Cristina Sampaio, Matthew B Stern, Richard Dodel, et al. Movement disorder society-sponsored revision of the unified parkinson's disease rating scale (mds-updrs): Scale presentation and clinimetric testing results. *Movement disorders*, 23(15):2129–2170, 2008.
8. John Staudenmayer, Shai He, Amanda Hickey, Jeffer Sasaki, and Patty Freedson. Methods to estimate aspects of physical activity and sedentary behavior from high-frequency wrist accelerometer measurements. *Journal of Applied Physiology*, 119(4):396–403, 2015.
9. Robert LeMoyne, Timothy Mastroianni, Michael Cozza, Cristian Coroian, and Warren Grundfest. Implementation of an iphone for characterizing parkinson's disease tremor through a wireless accelerometer application. In *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, pages 4954–4958. IEEE, 2010.
10. Avishai Wagner, Naama Fixler, and Yehezkel S Resheff. A wavelet-based approach to moniotring parkinson's disease symptoms. *International Conference on Acoustics, Speech, and Signal Processing*, 2017.
11. Rui Wang, Fanglin Chen, Zhenyu Chen, Tianxing Li, Gabriella Harari, Stefanie Tignor, Xia Zhou, Dror Ben-Zeev, and Andrew T Campbell. Studentlife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In *Proceedings of the 2014 ACM International*

Joint Conference on Pervasive and Ubiquitous Computing, pages 3–14. ACM, 2014.

12. Stanley R Kay, Abraham Flszbein, and Lewis A Opfer. The positive and negative syndrome scale (panss) for schizophrenia. *Schizophrenia bulletin*, 13(2):261, 1987.
13. Robert W Buchanan and Douglas W Heinrichs. The neurological evaluation scale (nes): a structured instrument for the assessment of neurological signs in schizophrenia. *Psychiatry research*, 27(3):335–350, 1989.
14. RS Wallwork, R Fortgang, R Hashimoto, DR Weinberger, and D Dickinson. Searching for a consensus five-factor model of the positive and negative syndrome scale for schizophrenia. *Schizophrenia research*, 137(1):246–250, 2012.

3.2 Topic Models for Automated Motor Analysis in Schizophrenia Patients

Accepted for publication in the *IEEE International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, 2018.

Topic Models for Automated Motor Analysis in Schizophrenia Patients

Talia Tron*[1,4] Yehezkel S. Resheff*[1,4] Mikhail Bazhmin[2] Abraham Peled[2,3] Alexander Grinsphoon[2,3] Daphna Weinshall[4]

***Abstract*— Wearable devices fitted with various sensors are increasingly being used for the automatic and continuous tracking and monitoring of patients. Only first steps have been taken in the field of psychiatric care, where long term tracking of patient behavior holds the promise to help practitioners to better understand both individual patients, and the disorders in general. In this paper we use topic models for unsupervised analysis of movement activity of schizophrenia patients in a closed ward setting. Results demonstrate that features computed on the basis of this analysis differentially characterize interesting sub-populations of schizophrenia patients. Positive-signs schizophrenia sub-population was found to have high *motor richness* and low *typicallity*, while negative-signs patients had low *motor richness* and lower *typicality*. In addition we design a classifier which correctly classified up to 80% of the clinical sub-population (f-score=0.774) based on motor features.**

I. INTRODUCTION

Motor peculiarities are an integral part of the schizophrenia disorder, both as aspects of the more general symptom repertoire, and in response to medications. To date, these symptoms are typically evaluated in a descriptive manner based on psychiatric rating scales such as the Positive and Negative Syndrome Scale (PANSS) [1], or targeted specifically using subjective clinical scales such as the Unified Dyskinesia Rating Scale (UDysRS) [2] and the Unified Parkinson's Disease Rating Scale (MDS-UPDRS) [3]. The lack of objective, quantitative methods for measuring these symptoms, and the insufficient conceptual clarity around them, may cause multiple interpretations of phenomenology, leading to low reliability and validity of diagnosis. In addition, the symptom evaluation process requires expert staff and availability of resources, and is therefore not done frequently enough to capture more subtle changes in spontaneous and drug-induced conditions. Clearly there is an urgent need for automatic monitoring and assessment tools.

The last decade has seen a steep rise in the use of wearable devices for medical applications in a range of fields, from human physiology [4] to movement disorders [5], [6] and mental health [7]. Accelerometers and gyroscopes, which are commonly embedded in smart-watches and other wearable devices, are now used to assess mobility and recognize activity. In a clinical setting, these sensors may be used in order to detect changes in high-level movement parameters, track their dynamics and correlate them with mental state.

Measures of activity such as step counts and overall activity, as well as changes thereof, have already been shown to effectively provide insights into the state of patients in a closed ward mental hospital setting [8]. Unsupervised behavioral mode analysis of sensor data, such as topic models, have previously been used in other domains to provide a high level description of behavior [9]. Here we combine these ideas and use topic models for unsupervised analysis of patient activity. These models allow a richer, qualitative description of behavior than the aforementioned measures. We demonstrate that features computed on the basis of topic model analysis differentiate sub-populations of patients.

II. MATERIALS AND METHODS

A. Study Design

27 inpatients from the closed wards at Shaar-Meashe MHC participated in the study. Most participants (21/27) were diagnosed with schizophrenia according to the DSM-5, 3 with paranoid schizophrenia, 2 with schizoaffective disorder, and one with psychotic state cannabinoids. Participants' age varied from 21 to 58 (mean 37.5), with course of illness varying from 0 (first hospitalization) up to 37 years (mean 16.9 years). Two of the patients dropped out of the study after less than a day due to lack of cooperation. The remaining 25 patients were followed for a period of three weeks on average (6-52 days).

The study was conducted in natural settings, where patients were *not* required to change any personal or medical procedure. On top of the normal care, every patient underwent an additional evaluation by a trained psychiatrist twice a week. The procedure included clinical evaluation of symptom severity using PANSS; Neurological Evaluation Scale (NES [10]) assessment was conducted as a control. In addition, continuous medication monitoring (type, dosage and frequency) by the clinical staff was observed.

All procedures performed in the study were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

B. Data Acquisition

Each participant was fitted with a smart-watch (GeneActiv[1]) with tri-axial accelerometer embedded sensors, the high

* T. Tron and Y.S. Resheff contributed equally to this work.
[1]The Edmond and Lily Safra center (ELSC) for Brain Science, Hebrew University, Jerusalem 91904, Israel talia.tron@mail.huji.ac.il
[2]Rappaport Faculty of Medicine, Technion Institute of Technology, Haifa 3200003, Israel
[3]Sha'ar Menashe Mental Health Center, Sha'ar Menashe 38706, Israel
[4]The Rachel and Selim Benin School of Computer Science and Engineering, Hebrew University, Jerusalem 91904, Israel

[1]https://www.geneactiv.org/

frequency output ($50Hz$) of which was stored on memory cards. Data was collected continuously throughout the experiment for a total of 489 days, from 25 patients. In order to reduce noise introduced by the variability in patient activity, the analysis focused on fixed time windows corresponding to regular departmental daily activity: Occupational therapy time slots (10am-11am), lunch (12pm-1pm), and indoor free time (4pm-5pm). In addition, we calculated full day features (6am-10pm) and used night time features (10pm-6am) to evaluate sleep quality. Weekends were excluded from the analysis.

III. ANALYSIS

A. Revising clinical assessment

The 30-item Positive and Negative Syndrome Scale (PANSS) was reduced to a five-factor description (Positive, Negative, Disorganized/Concrete, Excited and Depressed), according to the consensus model suggested by Wallwork et al. [11], based on 25 previously published models and refined with confirmatory factor analysis (CFA). Only the positive and negative factors were used for further analysis.

Clinical observations show that changes in a patient's symptoms occur continuously on a daily basis [12]. We therefore interpolated the bi-weekly PANSS factor scores, to achieve smooth daily scoring of symptoms. This was done using the PCHIP 1-d monotonic cubic interpolation, resulting in 494 data points (originally 118).

Interpolated data points were used to classify clinical sub-populations of patients on a daily basis. Sub-populations included patients with "High positive" symptoms, "High negative", "High negative and positive", and "Low" level symptoms. The remaining intermediate data points were discarded from the classification. This sub-typing allowed us to explore how different motor features are expressed in different clinical manifestations. Clustering was done based on the percentile of the positive and negative factors, each axis separately (Fig. 1).

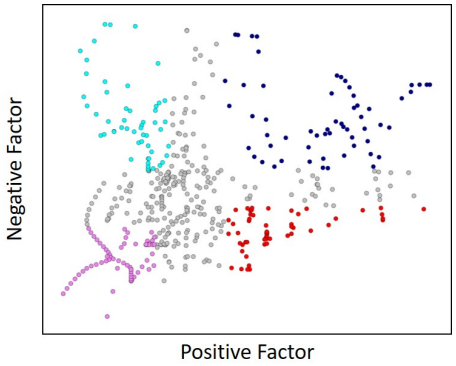


Fig. 1. Clinical sub-populations. Each data point represents the severity of the positive and negative factors for a specific patient in a specific day (N=494) based on the interpolated PANSS factors data. In the "Low" sub-population (magenta, N=65) both negative and positive symptoms lie in the bottom quartile, while in the "High negative and positive" (blue, N=59) both lie in the top quartile. The "High negative" sub-population (cyan, N=53) lies in the top vertical quartile with positive symptom values lower than median, while the "High positive" (red, N=57) lies in the top horizontal quartile with lower than median negative symptom values. The remaining data points (N=260) were classified as "Intermediate" (green).

B. Online computation of "patch features"

Topic model analysis requires the discretization of the continuous accelerometer signal both in time and in intensity, to produce word analogues – motor words. This mapping involves the creation of a code-book. The patch feature topic model procedure described in [9] contains a codebook generation stage where clustering (k-means) is applied to segments (a.k.a. patches) from the entire dataset. Given the larger dataset at hand, we designed an online greedy approximation to this procedure.

Specifically, the idea behind an online generation of code-book is to follow the way a dictionary would be created for a natural language corpus. The process proceeds with a single pass over the data. Each word is considered sequentially, and added to the dictionary on first encounter.

Since the words we are using describe the continuous accelerometer signal, we must also define what we mean by a word. Ideally, the dictionary should not be affected by small random changes in the signal. Additionally, since many behavioral modes are periodic to some extent, we would like the representation to allow wrap-around of patches. This would imply that the sequences of patches ABC and BCA, for example, have similar representation in the dictionary.

We achieve both these goals by using a discretized version of the signal and wrap-around equivalence classes. We use a SAX-like method [13] to encode each patch into a string. The process is as follows: Each interval on the time-axis is replaced by the mean value in the interval. Next, these point-values are replaced by a letter (discretized) according to their value. The output of this process is a string of length $\frac{patch-size}{interval-size}$ over a pre-determined alphabet.

Algorithm 1 Online codebook creation

```
codebook ← empty list
for each patch in the dataset do
    patch_word ← SAX_representation(patch)
    if patch_word (or equivalent) not in codebook then
        append patch_word to the codebook
    end if
end for
return codebook
```

The procedure resulted in 150K distinct words which described the entire dataset, distributed much as would be expected from a text corpus (see top panel in Fig. 2).

C. Topic Modeling over Motor Words

Latent Dirichlet Allocation (LDA) is a widely used topic model, with origins in natural language processing, and applications in many domains ranging from music modeling to motion of cars. On top of their traditional purpose of finding hidden semantic structures in data, these models have been shown to be useful for detecting surprising (or novel) events [14], [15].

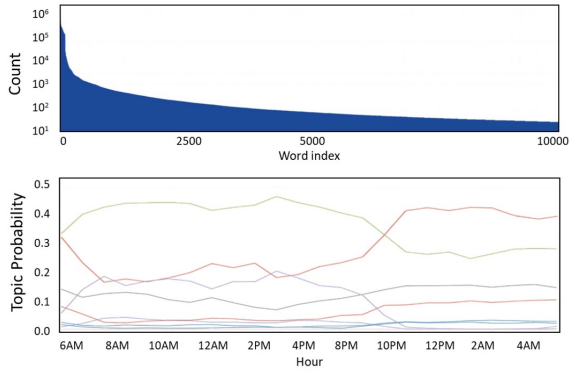


Fig. 2. Top - the motor-word frequency histogram for the entire dataset, truncated after the most common 10K words. These pseudo-words demonstrate the long-tail scale-free property characteristic of a natural language. Bottom- the daily topic distribution vector averaged over all patients.

Data was divided into blocks of 15 minutes of continuous signal; these serve as documents for the topic model, each represented as a histogram over the motor words as described above. The LDA process provides as output both a distribution over topics for each of the documents, and a distribution over words for each of the topics. Subsequently, a specific time window of a specific patient is characterized by a probability vector over the topics. The bottom panel in Fig. 2 demonstrates the distribution of the 10 topics used here over all patients and days. We can see that topic 6 (green) and topic 10 (purple) are typically prominent during the day, while other topics are more likely to occur during the night or throughout.

D. Topic Features

The advantage of using a data-driven unsupervised representation is that its features, unlike the supervised energy and step-count measures [8], are not directly connected with the intensity of the motor signal. Instead, this representation captures the *quality* of motor behavior in a given period of time. For example, a very repetitive behavior can be expressed by a low number of unique 'motor-words' in a specific time window. This allows us to compare patients behavior to themselves and to others in different activity windows, and thus be able to measure 'typicality' of the behavior for instance. Three Features were calculated based on topic models, separately for each data point (namely for each patient on each day, and each of the predefined activity windows described in section II-B):

1) Motor Richness: The normalized distinct word count per activity window.

$$\text{Motor Richness} = \frac{\text{distinct word count in window}}{\text{window length}}$$

This measure represents the range of motor activity repertoire. A low score implies that the patient repeatedly performed similar movements, while a high score corresponds to the use of many different movement patterns.

2) Consistency:

$$\text{Consistency} = 1 - D_{KL}(v \parallel \bar{v})$$

where v denotes the topic distribution over the time window, and $\bar{v}$ the mean topic distribution for the patient in the same window over all measured days. D_{KL} denotes the Kullback-Leibler divergence. This score measures how regular the patient's motor behavior is in the given time window.

3) Typicality: the entropy of the topic distribution vector.

$$\text{Typicality} = H[v] = -\sum_{i=1}^{10} v_i \log(v_i)$$

Low entropy implies that a small number of topics can capture the activity. High entropy implies that the observed activity is a mixture of many topics. We name this measure *typicality* since typical activity should be captured by one (or a few) topics, thus producing low-entropy topic distributions [15].

The manifestation of each feature in clinical sub-populations was tested using one-way ANOVA separately for each of the activity windows. In addition, a learning algorithm for automated sub-population classification was designed and evaluated.

E. Classification Algorithm

Classification was carried out using a two-step algorithm based on linear support vector machines (SVM) and decision trees classifiers, in order to distinguish between different sub-populations (described in III-A) based on motor features (Algorithm 2). The algorithm was trained to discriminate sub-populations, and specifically classify "High positive" vs. "High negative" and "Low" vs. "High positive and negative".

In the first step, individual classifiers were trained for each activity window separately (lunch, occupational therapy, free-time, day, night, and all). In the second step, the probabilistic output of the 6 time-specific first-stage classifiers was used to train a second, daily-model, which determined the clinical category.

Feature selection was done based on the ANOVA f-values of each individual feature on train data. These were calculated separately for each activity window, and the same features were used also for testing.

Algorithm 2 Two stage algorithm for patient sub-type classification based on activity in time-windows.

1: **for all** time-windows w_i **do**
2: train base classifier c_i on w_i and target y
3: **end for**
4: **for all** time-windows w_i **do**
5: $\hat{y}_i \leftarrow$ prediction of c_i on w_i
6: **end for**
7: train final classifier c on the set of first-stage predictions $\hat{y}_i$ and target y

The algorithm was evaluated in a leave-one-out framework, where in each iteration a different observation (specific

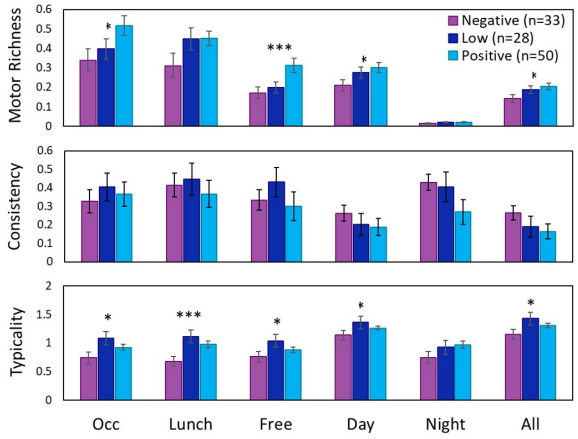


Fig. 3. ANOVA results for topic features in 3 clinical sub-populations (see Fig. 1): "High positive" (denoted *Positive*), "High negative" (denoted *Negative*), and "Low". The analysis was repeated separately for each time window (X-axis). *Motor richness* was highest in the *Positive* sub-population (cyan) and lowest in the *Negative* sub-population (magenta). This was most significant during free time, but was also true for all other activity windows (p-values between 0.05-0.07 are marked by half an asterisk). *Typicality* was generally highest in the *Low* sub-population, and lowest in the *Negative* sub-population, with the most significant difference during lunch time. No significant group different was found for *Consistency* although it was lowest in the *Positive* sub-population in all activity windows.

patient in a specific day) was left out and the model was trained on the remaining data and tested on the left out sample. To avoid possible contamination of test data (leakage) due to observation interpolation, when using an interpolated point as the test, all actual observations it was based upon were excluded from the train data.

IV. RESULTS

A. Motor Activity in different Clinical Sub-populations

Fig. 3 summarizes the results of subjecting all features to ANOVA analysis. *Motor richness* is consistently highest for the "High positive" sub-population, and lowest for the "High negative" sub-population, with the "Low" sub-population somewhere in the middle. This indicates that patients with active positive symptoms tend to have a higher variety of motor activities, while negative symptoms are expressed in poorer movement repertoire. The trend was evident in all activity windows but was only found significant during free time ($F = 5.09$, $p = 0.0077$).

As expected, *typicality* is highest for the "Low" sub-population, consistently over all activity windows. The lowest *typicality* is observed in the "High negative" sub-population, indicating that the motor activity of these patients is less similar to the common motor behavior. The biggest group difference was found over lunch time ($F = 7.48$, $p = 0.00090$) but it was also significant during occupational therapy ($F = 4.39$, $p = 0.015$), free time ($F = 3.38$, $p = 0.037$) and throughout the day ($F = 3.78$, $p = 0.026$). No group difference was found for *consistency*, although it was lower in the "High positive" sub-population in all activity windows.

B. Classification Results

For "High positive" vs. "High negative" classification, the best results were achieved using linear SVM for the first stage (window-based model, see Algorithm 2) with top-5 selected features, and decision tree for the second (daily) stage. The algorithm correctly classified 78% of the"High negative" observations, and 58% of the "High positive" observations. All together the mean precision was 0.651 and mean recall was 0.654 on test data (f-score=0.652).

Slightly better results were achieved for the "Low" vs. "High positive and negative" classification, using linear SVM for both stages and top-5 selected features. Here the algorithm correctly classified 81% of the "Low" observations and 70% of the "High positive" observations. All together the mean precision was 0.757 and mean recall was 0.748 on test data (f-score=0.774).

REFERENCES

[1] S. R. Kay, A. Flszbein, and L. A. Opfer, "The positive and negative syndrome scale (panss) for schizophrenia.," *Schizo. bulletin*, vol. 13, no. 2, p. 261, 1987.

[2] C. G. Goetz, J. G. Nutt, and G. T. Stebbins, "The unified dyskinesia rating scale: presentation and clinimetric profile," *Mov. Dis.*, vol. 23, no. 16, pp. 2398–2403, 2008.

[3] C. G. Goetz, B. C. Tilley, S. R. Shaftman, G. Stebbins, *et al.*, "Movement disorder society-sponsored revision of the unified parkinson's disease rating scale (mds-updrs): Scale presentation and clinimetric testing results," *Mov. dis.*, vol. 23, no. 15, pp. 2129–2170, 2008.

[4] J. Staudenmayer, S. He, A. Hickey, J. Sasaki, and P. Freedson, "Methods to estimate aspects of physical activity and sedentary behavior from high-frequency wrist accelerometer measurements," *Journal of Applied Physiology*, vol. 119, no. 4, pp. 396–403, 2015.

[5] R. LeMoyne, T. Mastroianni, M. Cozza, C. Coroian, and W. Grundfest, "Implementation of an iphone for characterizing parkinson's disease tremor through a wireless accelerometer application," in *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, pp. 4954–4958, IEEE, 2010.

[6] A. Wagner, N. Fixler, and Y. S. Resheff, "A wavelet-based approach to moniotring parkinson's disease symptoms," *International Conference on Acoustics, Speech, and Signal Processing*, 2017.

[7] R. Wang, F. Chen, Z. Chen, T. Li, G. Harari, S. Tignor, X. Zhou, D. Ben-Zeev, and A. T. Campbell, "Studentlife: assessing mental health, academic performance and behavioral trends of college students using smartphones," in *Proceedings of the 2014 ACM Int Joint Conf on Pervasive and Ubiquitous Computing*, pp. 3–14, ACM, 2014.

[8] T. Tron, Y. S. Resheff, M. Bazhmin, A. Peled, and D. Weinshall, "Real-time schizophrenia monitoring using wearable motion sensitive devices," in *MobiHealth*, Springer, 2016.

[9] Y. S. Resheff, S. Rotics, R. Nathan, and D. Weinshall, "Topic modeling of behavioral modes using sensor data," *International Journal of Data Science and Analytics*, vol. 1, no. 1, pp. 51–60, 2016.

[10] R. W. Buchanan and D. W. Heinrichs, "The neurological evaluation scale (nes): a structured instrument for the assessment of neurological signs in schizophrenia," *Psychiatry research*, vol. 27, no. 3, pp. 335–350, 1989.

[11] R. Wallwork, R. Fortgang, R. Hashimoto, D. Weinberger, and D. Dickinson, "Searching for a consensus five-factor model of the positive and negative syndrome scale for schizophrenia," *Schizophrenia research*, vol. 137, no. 1, pp. 246–250, 2012.

[12] S. Arieti, *Interpretation of schizophrenia*. Basic Books (AZ), 1974.

[13] J. Lin, E. Keogh, S. Lonardi, and B. Chiu, "A symbolic representation of time series, with implications for streaming algorithms," in *Proceedings of the 8th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery*, pp. 2–11, ACM, 2003.

[14] U. Shalit, D. Weinshall, and G. Chechik, "Modeling musical influence with topic models," in *International Conference on Machine Learning*, pp. 244–252, 2013.

[15] A. Hendel, D. Weinshall, and S. Peleg, "Identifying surprising events in videos using bayesian topic models," in *Asian Conference on Computer Vision*, pp. 448–459, Springer, 2010.

3.3 ARIMA-based motor anomaly detection in schizophrenia inpatients

Accepted for publication in the *IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)* 2018.

ARIMA-based Motor Anomaly Detection in Schizophrenia Inpatients

Talia Tron[1,4] Yehezkel S. Resheff[1,4] Mikhail Bazhmin[2] Daphna Weinshall[4] Abraham Peled[2,3]

Abstract— **Motor alteration is an important aspect of the elusive schizophrenia disorder, manifested both throughout the various phases of the disease and as a response to treatment. Tracking of patients' movement, and especially in a closed ward hospital setting, can therefore shed light on the dynamics of the disease, and help alert staff to possible deterioration and adverse effects of medication. In this paper we describe the use of ARIMA-based anomaly detection for monitoring of patient motor activity in a closed ward hospital setting. We demonstrate the utility of the approach in several intriguing case studies.**

I. INTRODUCTION

Monitoring of motor behavior is part of the regular assessment of schizophrenia patients and is vital to diagnosis, progress assessment and to the monitoring of medication response. Various alterations of motor behavior are evident throughout the phases of the disease, and as a response to treatment. The psychotic acute phase of schizophrenia is typically accompanied by restlessness, including occasional bizarre movements and gestures, while post psychotic deficiency negative symptoms are related to reduced activity, slowness and even freezing. Antypsychotic medications may cause Parkinsonism, i.e., tremor, rigidity, and slowness, which usually pass after the first week of treatment.

Despite its clinical and diagnostic value, to date, motor monitoring is done in a descriptive non etiological manner based on subjective clinical scales, which may result in biased, inaccurate and typically non quantifiable assessments. This kind of assessment requires expert staff and the availability of resources, and may not be frequent enough to capture significant changes in spontaneous and drug-induced conditions. These issues can be alleviated by carrying out objective, continuous quantifiable monitoring [1], the investigation of which is the goal of this study. Accelerometers and gyroscopes, commonly embedded in smart-watches and other wearable devices, have been extensively used over the last decades in medical applications ranging from human physiology [2] to movement disorders [3] and mental healthcare [4]. These cheap and widely available sensors may be used for continuous qualitative patient monitoring in natural clinical settings. Accelerometer data have already been shown to effectively provide insights into patients clinical state, and motor features were successfully used for clinical sub-typing in a closed ward mental hospital setting [5], [6]. Here we focus on detecting acute abnormal which are either the result or the cause of drug modifications or changes in patients' clinical conditions. Our approach employs forecasting models widely used in statistics and econometrics, applied to step-count data. We demonstrate the utility of this approach with 4 schizophrenia case studies, in which we evaluate monitoring performance based on medical and clinical records.

[1]The Edmond and Lily Safra center for Brain Science, Hebrew University, Jerusalem 91904, Israel talia.tron@mail.huji.ac.il

[2]Rappaport Faculty of Medicine, Technion Institute of Technology, Haifa 3200003, Israel

[3]Sha'ar Menashe Mental Health Center, Sha'ar Menashe 38706, Israel

[4]The Rachel and Selim Benin School of Computer Science and Engineering, Hebrew University, Jerusalem 91904, Israel

II. MATERIALS AND METHODS

A. Study Design

Four inpatients from the closed ward at Shaar-Menashe mental health center, diagnosed with schizophrenia according to the DSM-5, participated in the study. One patient (patient B) was diagnosed with paranoid schizophrenia. Participants' age varied from 24 to 54 (average 36.9), with course of illness varying from 7 to 35 years (average of 13.5 years). After signing the appropriate Helsinki legal consents, participants were tracked for a period of approximately one month (27-31 days) in natural settings. During this period, patients were monitored for medication use (type, dosage, and frequency) by the nurses and the physicians. In addition, every patient underwent a clinical evaluation of Positive and Negative Syndrome Scale (PANSS [7]) and Neurological Evaluation Scale (NES [8]) by a trained psychiatrist twice a week. The neurological evaluation was only utilized to confirm that no psycho-motor deficits were evident in any of the participants during the experiment.

All procedures performed in the study were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

B. Data Acquisition

At study onset, participants were given smart-watches with embedded accelerometers (GeneActiv[1]). These watches were worn on the wrist throughout the experiment. The output ($50Hz$) of the sensors was stored on internal memory cards. The study was conducted in natural settings, where patients were *not* required to change any personal or medical procedure. None of the patients expressed any discomfort or disturbance from wearing the device.

III. DATA ANALYSIS

A. Building personal ARIMA Models

Analysis focused on the walking pattern of patients, aiming to detect significant quantitative changes. Stepping

[1]https://www.activinsights.com/products/geneactiv/

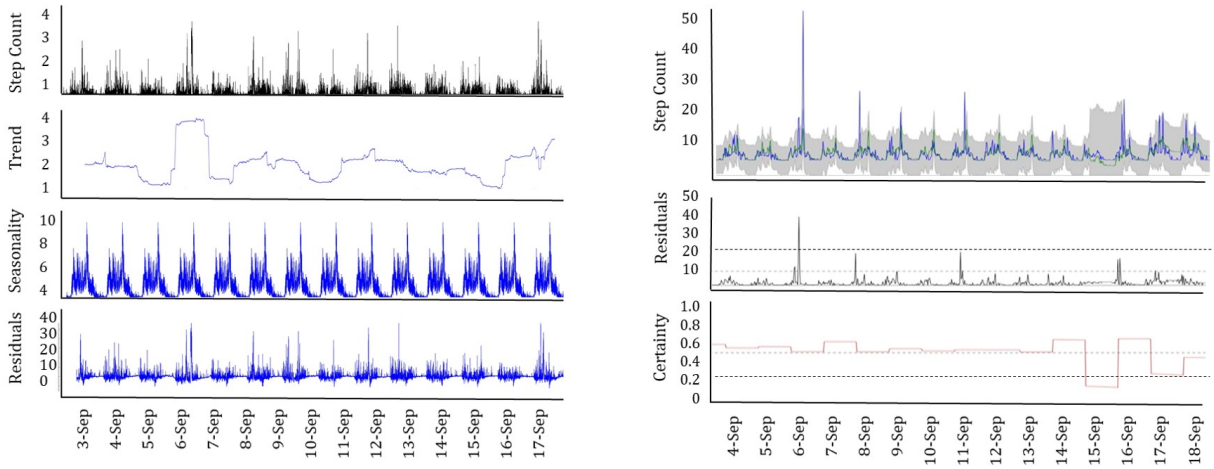


Fig. 1. **Left**- Decomposition of daily steps (top) of a single patient to trend (smoothed series calculated using centered moving average), seasonality (regularly repeating data patterns calculated as the average of the smoothed series for each period) and noise. **Right**- Demonstration of the ARIMA model for patient A. The model returns the predicted mean and a 95% confidence interval (CI) around it. Abnormal behavior is detected when (a) the observed step count value lies outside the CI predicted by the model, (b) the residuals are higher than threshold (e.g. September 6), or (c) when certainty is lower than threshold (e.g. September 15).

behavior was detected as large maxima of the smoothed square norm of the 3-axial $50Hz$ point-wise acceleration, and the number of steps (step count) was averaged over 1 minute intervals (see [5] for further details).

We used AutoRegressive Integrated Moving Average (ARIMA) models to detect abnormal walking patterns. One week of data was used to predict the step count for the following day, together with the associated confidence interval. Repeating this in a rolling window design produced the predicted step count for the entire duration of available data, around 3 weeks for each participant excluding the first week. Predicted values were then compared to those observed in practice for the purpose of anomaly detection [9].

We began by decomposing the step-count data into trend, seasonality and noise components, as shown in the left side of Fig. 1. As expected, strong daily seasonality was seen in the data. It is interesting to note that the trend component, to the extent that it exists, may potentially be used for direct real-time monitoring of patients symptom severity over time.

Next, we aggregated each patient's step-count data in windows of 10-30 minutes (this was done to smooth the data on the one hand, and reduce computation on the other). Both regular and daily seasonal differentiation were computed to obtain a stationary signal. We applied 4 different ARIMA models to all patients, and evaluated them using AIC criteria with mean and absolute errors. The emerging preferred model was ARIMAX(1,1,1) seasonal (1,1,2), which had a consistent lower error and lower AIC over all patients.

B. Abnormal behavior detection

For each patient separately, we ran an ARIMAX(1,1,1) seasonal (1,1,2) model, which was based on 7 days of data in order to predict the following day. The model provided the predicted mean and a 95% confidence interval (CI) around it. Model residuals were calculated as the squared difference between the model predicted values and the observed values during the test period.

A measure of prediction certainty was calculated based on the normalized CI size ($|CI_z|$) as follow:

$$Certainty = 0.95 \times 2 \times \frac{std(data)}{|CI_z|} \tag{1}$$

This is a measure of model confidence, with low values indicating that the model hasn't been able to accurately predict future values based on the patient's history. The multiplier of 0.95 sets the maximum certainty value to 0.95 (model confidence level). Although certainty is somewhat correlated with residuals size, this is an important independent measure. Specifically, it covers cases where the observed value is lower than the predicted value, which is not always expressed in CI range or high residuals.

Abnormal behavior is defined as one the following (see right side of Fig. 1): (a) The predicted value is not in the model CI; (b) the residuals between model prediction and observed values are higher than threshold (set to be 3 times the mean residuals on train data); (c) the $certainty$ of the model is lower than threshold (0.3). In order to avoid trailing errors and secure robustness, when abnormal behavior is detected, the observed values of the training period are replaced with predicted values. On repeated detections (more than twice) the model is adjusted back to observed values.

C. Evaluating model performance

In order to evaluate our model we systematically studied the patients clinical records and drug charts, and compared them with model anomaly detections. No clear abnormal event, such as an outburst of violence or riot, was recorded during the experiment period. We therefore used the PANSS clinical records in order to identify *abnormal* events, which are time stamps corresponding with a steep increase or

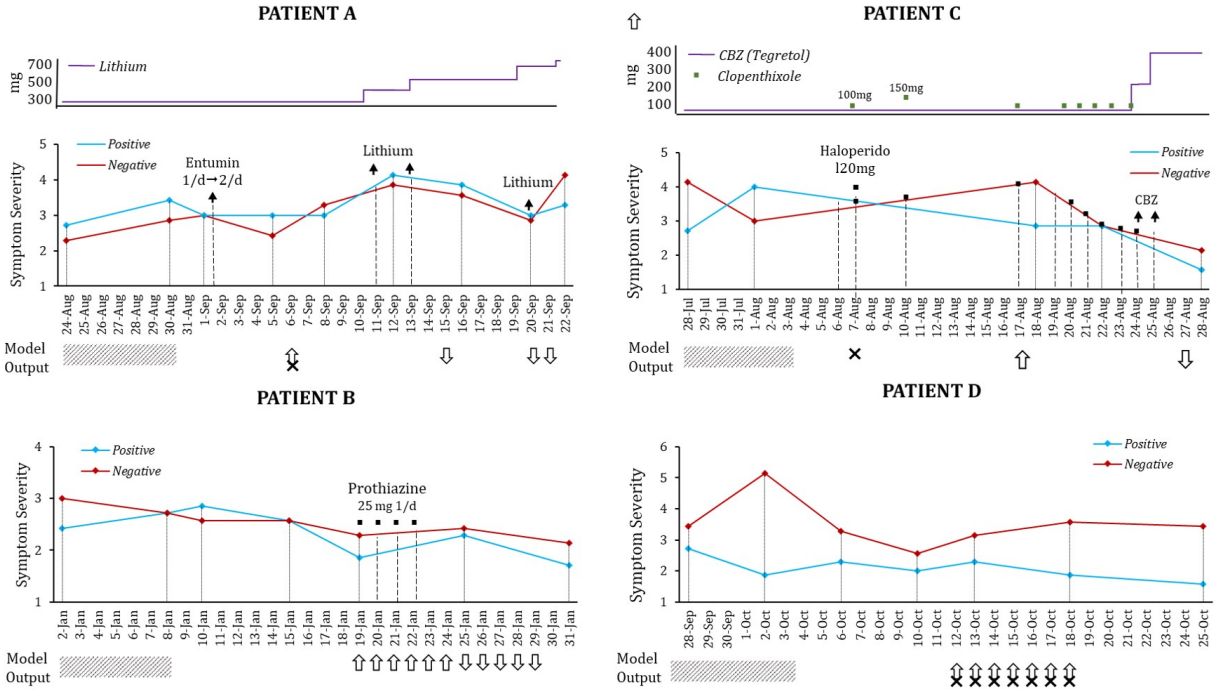


Fig. 2. Description of model prediction vs. clinical and medication records monitoring for all four patients. The direction of the white arrows in the bottom part of each graph indicates whether increased activity (up) or decreased activity (down) has been detected. A cross under the arrow indicates unexplained detection, while a cross without an arrow indicates an event that wasn't detected by the model. The dashed rectangle marks the training period of the model. In the line chart above, the mean severity of positive (red) and negative (blue) symptoms is shown. The black symbols indicate a change in drug dosage (arrow) or a single administration (square). In case of dosage change, the top graph (in patients A and C) indicates its amount (in mg).

decrease in symptom severity (more than 2 degrees on the PANSS scale) between two clinical sessions. Results are summarized in Fig. 2.

In an effort to capture some larger scale dynamics, we took note of the general positive and negative symptoms trend. Every change in drug dosage was also considered an abnormal event, since these changes are rare and usually indicate a change in a patient's clinical condition. It should be noted that increased drug dosage may be either a response to abnormal activity (when the detected event took place prior to drug adjustment) or its trigger (when the detected event followed a drug adjustment). Decreased dosage, on the other hand, is usually followed by continuous improvement in symptom severity, but may still cause side effects. Therefore, in order to obtain a coherent picture, both timing and the direction of the dosage change were taken into account.

For each abnormal event detected by our model, we looked for an explanation (as defined above) in the clinical records (drug dosage and PANSS scores); an event which did not have a satisfactory explanation, was labeled as 'unexplained'. Likewise, a drug change event or a steep change in the clinical evaluation data which was not detected by our model was labeled as 'undetected'. The number of unexplained and undetected events was used to roughly estimate the accuracy and sensitivity of our model. Events in consecutive days were counted as one continuous event.

1) Patient A: Abnormal increased walking behavior was detected on September 6th. On the same day, the dosage of *entumin* (a.k.a *clotiapine*), an atypical anti-psychotic drug, was increased from 40mg 1/day to 40mg 2/day.

On September 15th, and then again during September 20-22, our model detected lower than expected activity. In the clinical records, we see a significant increase in both positive and negative symptoms during September 5-12, with a steep rise in active social avoidance, hostility and social withdrawal. Possibly this behavioral change has resulted from the increased entumin dosage, although we cannot rule out other possible triggers.

Following this deterioration in the patient's condition, on September 11th the dosage of *lithium* was increased, and again on the 13th. Both positive and negative symptoms were reduced in subsequent days, with active social avoidance and hostility returning to normal values. We also see the emergence of increased negative symptoms, including blunted affect and passive apathetic social withdrawal.

Lithium is known to take effect within 1-3 weeks, so the lower activity found by our model during September 20-22 may be the result of the September 11th dosage increase. The September 15th detection remains unexplained by drug records but is congruent with clinical data.

In summary, 2/3 detected events for this patient had a co-found explanation in the clinical and medication records. One event had only a weak co-found in the clinical data. No clinical trend or drug changes remained undetected.

2) Patient B: The model detected a period of extreme increased activity during January 19-24, followed by decreased

activity during January 25-31. On January 19th, this patient was given *prothiazine*, a neuroleptic medication used as a sedative and weak anti-psychotic, for a period of 4 days. We found no significant change in symptom severity for this patient prior to the sedative drug administration, with only a small decrease in overall negative symptoms at that time. This is probably because clinical evaluation was not frequent enough to capture the change. The fact that our model detected this event while the clinical data did not, can be used as evidence for the potential benefit of continuous automated monitoring.

On January 25th, two days after the patient has stopped receiving the medication, we see a small improvement in his clinical condition with normal level of motor activity. In the model this is expressed by a detected 'lower than expected' activity, based on the increased activity in the previous days.

In summary, for this patient all detected events (2) had a co-found explanation in the medication records but no co-found (or a minor one) in the clinical records. No clinical trend or medication alteration remained undetected.

3) Patient C: Increased activity level was detected by the model on August 17th. Clinical data together with medical records clearly suggest that around this period there was an aggravation in the patient's condition. On August 17th, he was injected with 100mg of *clopenthixole acetate* (anti-psychotic and acute sedative medication), and once again in the following days (August 20-25). The drug's effect seems to have been dimmed unsatisfactory, since during August 24-25 the patient was also prescribed 200mg and then 400mg of *carbamazepine* (CBZ), an off label medication used in combination with anti-psychotics when the treatment with anti-psychotics alone has failed [10]. In the clinical data we see a decrease in both negative and positive symptoms severity around August 18-22, with a steep decrease in hallucinations, poor attention, and motor retardation. This improvement is most probably the result of the massive drug treatment. On August 27th, after the patients symptoms were reduced and drug treatment was stabilized, the model detected a significant reduction in patient's activity.

In August 7 the patient received two types of typical anti-psychotic medications (clopenthixole and *haloperidol*), and then again in August 10 (only *clopenthixole*). Since these drugs act on a short term basis, it is not probable that the the worsening in the patient's condition in subsequent days was triggered by this medication change. The most probable explanation is that there was some acute event at that time, which was not detected by our model.

In summary, all detected events (2) had a co-found explanation in the clinical and medication records, while one likely clinical event remained undetected.

4) Patient D: The model reported a period of decreased activity during October 12-18, with low certainty. No medication change was registered in this time period, and no substantial evidence was found in the clinical data (only a steep increase in stereotyped thinking). The overall trend of symptoms' change around that period leaned towards increased negative symptoms and reduced positive symptoms. This happened following approximately a week of steep decrease in negative symptoms.

In summary, the event detected by our model had no co-found explanation in the medication records. No clinical trend or medication alteration remained undetected.

TABLE I
SUMMARY OF ANOMALY DETECTION RESULTS AND PATIENTS' DATA.

	Days	Sessions	Explained	Missed
Patient A	31	10	2/3	0
Patient B	29	7	2/2	0
Patient C	31	11	2/2	1
Patient D	27	7	0/2	0

As summarized in Table I, when aggregating data from all patients, 6/8 anomaly events detected by our model had a co-found explanation in the medication and clinical records (precision of 75%). 6/7 events were detected by our model, with one certain mis-detection in patient C (recall of 85%).Other detected events may have alternative explanation not available to our experimental design.

IV. CONCLUSIONS

Our study demonstrates the benefits of using forecasting models in conjunction with accelerometer data for the continuous monitoring of schizophrenia patients. In three out of four case studies, we found a direct link between detected behavioral events and changes in the patient's clinical condition or drug regime.

REFERENCES

[1] Christopher G Goetz, John G Nutt, and Glenn T Stebbins. The unified dyskinesia rating scale: presentation and clinimetric profile. *Mov. Dis.*, 23(16):2398–2403, 2008.

[2] John Staudenmayer, Shai He, Amanda Hickey, Jeffer Sasaki, and Patty Freedson. Methods to estimate aspects of physical activity and sedentary behavior from high-frequency wrist accelerometer measurements. *Journal of Applied Physiology*, 119(4):396–403, 2015.

[3] Avishai Wagner, Naama Fixler, and Yehezkel S Resheff. A wavelet-based approach to moniotring parkinson's disease symptoms. *International Conference on Acoustics, Speech, and Signal Processing*, 2017.

[4] Rui Wang, Fanglin Chen, et al. Studentlife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In *Proceedings of the 2014 ACM Int Joint Conf on Pervasive and Ubiquitous Computing*, pages 3–14. ACM, 2014.

[5] Talia Tron, Yehezkel S Resheff, Mikhail Bazhmin, Abraham Peled, and Daphna Weinshall. Real-time schizophrenia monitoring using wearable motion sensitive devices. 2016.

[6] Talia Tron, Yehezkel S Resheff, Mikhail Bazhmin, Alexander Grinshpoon, Abraham Peled, and Daphna Weinshall. Topic models for automated motor analysis in schizophrenia patients*. In *Body sensor networks (BSN), 2013 IEEE international conference on*. IEEE, 2017.

[7] Stanley R Kay, Abraham Flszbein, and Lewis A Opfer. The positive and negative syndrome scale (panss) for schizophrenia. *Schizo. bulletin*, 13(2):261, 1987.

[8] Robert W Buchanan and Douglas W Heinrichs. The neurological evaluation scale (nes): a structured instrument for the assessment of neurological signs in schizophrenia. *Psychiatry research*, 27(3):335–350, 1989.

[9] Varun Chandola, Arindam Banerjee, and Vipin Kumar. Anomaly detection: A survey. *ACM computing surveys (CSUR)*, 41(3):15, 2009.

[10] Daniela Ceron-Litvoc, Bernardo Garcia Soares, Geddes, et al. Comparison of carbamazepine and lithium in treatment of bipolar disorder: a systematic review of randomized controlled trials. *Human Psychopharmacology: Clinical and Experimental*, 24(1):19–28, 2009.

Chapter 4

Discussion and Conclusions

This dissertation presents the work done on automatic analysis of nonverbal behavior in schizophrenia patients. We offer innovative algorithmic and conceptual approaches for analyzing this behavior in experimental and natural settings. Our results demonstrated how the use of assistive technologies for facial expression and motor behavior analysis can contribute to the important process of diagnosis and monitoring in schizophrenia.

Although facial alterations and motor deficits have a clear diagnostic value in schizophrenia, not many studies has attempted to assess these behaviors in an objective automated manner. In the ones that do so, behavioral analysis is often restricted to general non ecologically relevant measurements, leading to a substantial lack of applicable tools and knowledge in the field. Our study objectives were thus as follows. First, we aimed to obtain impartial, accurate, informative measurements of patients behavior to be used for automated patients evaluation and monitoring. Second, we used these measurements in order to refine the definition of behavioral clinical concepts and to gain important insights regarding patients behavior.

We did so by focusing on automatic analysis of facial and motor activity, each with its own advantages and disadvantages. While facial expression data is very rich and informative, it is typically more difficult to obtain and process. Placing a camera in front of a patient requires cooperation, and can only be done under controlled conditions. In addition, facial tracking technology is still immature, and the leading tools in the field require individual training to obtain a user-specific models. Other solutions are based on generative facial models which are less precise, and sometime sensitive to user location and head rotation.

Actigraphy data on the other hand, has the clear advantages of being easily recorded using embedded sensors which can be worn continuously by patients. This allows one to gain information regarding subtle motor changes over time and in natural settings. Nevertheless, the data is very noisy and

not as informative as facial data regarding user's emotional and mental state. The commonly derived measures are of general motor activity level (number of steps, amount and variability of movement) where for many usages and application, in particular in the medical field, more precise and descriptive motor behavioral measures are required. Deriving such measures and recognizing specific activities (e.g. teeth brushing, stairs climbings) require prior knowledge about signal structure or about behavioral context. Otherwise, great amount of labeled data is needed to evaluate activity detection and classification algorithms.

In addition to the inherent technological limitation of non-verbal measures extraction, the interpretation of such measures may be subjected to theoretical misconceptions and biased. The most prominent example is the widespread belief that facial expressions signal people's emotion, which is not always the case in schizophrenia patients. Also, the categorical division of facial activity to a set of specific prototype emotions such as anger, sadness and disgust may be an over simplified and imprecise, as discussed in detail in the introduction (see 1.1.1).

Main Innovations

In order to achieve our research objectives, overcome the many technological limitations and avoid possible biases, we explored in this work new algorithmic and theoretical approaches.

First, we expanded the commonly used activity *level* descriptors of non-verbal behavior to a broader set of measures including variability, dynamics, consistency and appropriateness. Dynamic measures describe the amount and pace of change in non-verbal activity, and were calculated directly as the variance of activity level descriptors over time. More specific dynamic measures were often calculated in facial data using transition matrix representation, counting the number of changes and their extent, as described in 2.1. For motor data, steps counts can be seen as a dynamic measure, since it averages the number of steps over a period of one minute, somewhat correlated with walking velocity, although not necessarily. The measures of intensity and dynamics, seem to hold a great diagnostic and clinical promise. First, they are independent of emotional value, and are therefore less subjected to personal interpretation. In addition, they have the potential to be measured in a precise continuous manner, enabling clinicians to detect subtle changes

in patients' behavior as well as general trends over time. Lastly, these measures may be directly associated with the motor aspect of the disorder, as elaborated below.

Variability measures assess the diversity of non-verbal activity, beyond mere intensity (more or less activity), and were calculated on our data using unsupervised learning methods. Such measures include the facial-clusters features described in 2.2, obtained by k-means clustering over all AUs. For motor data, we discretized both the time and the intensity level of the accelerometer data in order to generate a codebook of 'motor words'. This in turn was used to obtain a topic-model based representation of motor richness and consistency as described in 3.2. These variability measures allow for a more qualitative description of facial and motor behavior without any prior assumptions, making it easier to compare between different clinical populations and within the patients over time.

In our study we derived highly relevant measures of non-verbal consistency and appropriateness, emphasizing the importance of taking into account behavioral context. In order to obtain such measures for facial activity, we used the well documented photo rating paradigm of the *International Affective Picture System* (IAPS) database. This experimental paradigm gave us information regarding patients emotional state (how did they feel while watching the photos) as well as their emotional congruity (how similar were their emotions to those occurring in the healthy population). However, it was time consuming, and did not necessarily reflect patients facial response in natural conditions. As for motor activity, here behavioral context was derived using prior knowledge regarding the patients' routine daily activities. This method gives a good balance between the restrictions of applying predesigned experimental paradigm and the noise introduced in natural studies. It allowed us to compare patients behavior over days and retrieve consistency and typicality measures. Another example of using contextual data can be found in chapter 3.3, where we used seasonal autoregressive models which take into account the 24 hours cycles in patients activity. Averaging over these cycles, strengthens the behavioral signal against the overall noise and considerably improved abnormal activity detection results.

Our work also offers some important novelties in the analysis of clinical data. Similarly to other data-driven studies, facial and motor activity measures were used as features for automated clinical classification and prediction (adjusted versions of support vector machine (SVM), decision tree and regularized regression were employed). However, instead of focusing

merely on group comparison (patients vs. controls), we identified and characterized clinical sub-populations. To our knowledge this is the first attempt to separate positive-signs and negative-signs schizophrenia, characterize each phase, and automatically discriminate between them. This sub-typing was enabled due to some clinical prior knowledge. Fore example, the 5-factor decomposition of PANSS clinical evaluation was based on literature review, and the interpolation of patients' clinical evaluation in 3.2 was based on the clinical observation that changes in a patient's symptoms occur continuously on a daily basis.

An additional novelty is using accelerometer data for clinical anomaly detection as done in 3.3. We demonstrated how time forecasting models can be applied to detect possible deterioration in single patient's behavior and detect adverse effects of medication.

Finally our study offers some important conceptual innovations with added diagnostic value when directly addressing the commonly used, yet often imprecisely defined concepts of *Affective Flatness* and *Affective incongruence*, and outpointing a possible confound.

Although our study focused on schizophrenia, the innovations it offers can be implemented for other mental disorders such as anxiety, depression and bi-polar disorder, where affect and movement are known to play a major role. In addition, some of the concepts and tools we present here are of high relevance in the healthy population, and may contribute to the affective computing (AC) and social signal processing (SSP) communities.

Results overview and discussion

The results of the facial activity experiments, suggest that *Affective flatness* is expressed not only as a reduction in the intensity of facial expressions, but also as a slowdown in facial dynamics and a restriction of expressional variability. In addition, our results point out a possible confusion between flatness and inappropriateness, where the same facial behavior may be interpreted as both, depending on the clinical context. When quantitatively taking into account overall flatness, patients' facial response to emotional stimuli did not significantly differ from that of the healthy population, rather it was somewhat less consistent. Namely, similar feelings were expressed in a wider range of relative facial behaviors.

This observation should highlight the unique role facial expressions play in schizophrenia and should be taken into consideration in the clinic. Whereas

in healthy individuals we are used to referring to the face as reflecting the mood, in schizophrenia this coupling seems to be disturbed. Our results reinforces the 'inhibition theory', which claims that patients with schizophrenia suffer from the inability to express the emotion they experience. Accordingly, we did not find any evidence for impaired emotional experience in patients. Namely, although facial activity in patients is less expressive, it does not necessarily reflect any emotional deficit.

In the context of motor behavior, we demonstrate how various motor alteration are manifested in different phases of the schizophrenia disorder. Consistent with previous studies and clinical observations, we found *positive-signs schizophrenia* to be associated with higher level of motor activity (step counts, total energy), and vice-versa for *negative-signs schizophrenia*. Our work further extended the discussion to more qualitative motor related features. We found that the motor repertoire of *positive-signs schizophrenia* patients is richer than that of controls, while *negative-signs schizophrenia patients* demonstrate a poorer, non-diverse repertoire. *Negative-signs schizophrenia* was also characterized by low typicality, namely their behavior was less similar to that of the general inpatients population.

The machine learning algorithms we designed correctly classified patients clinical condition with accuracy of over 80% (AUC=0.9) based on facial features (for patients vs. controls discrimination) and up to 78% using motor data (for clinical sub-population). The predictions were also in good correlation with symptom severity assessments. In addition, the forecasting seasonal models we developed used for personal patient's monitoring and abnormal behavior detection were demonstrated to be feasible on 4 patients case studies.

A few caveats regarding the current study should be kept in mind. First, most of the participating patients in the facial activity experiments had only negative signs of schizophrenia, and were receiving anti-psychotic medications. In the motor behavior experiment on the other hand, all participants were inpatients and comparison was made between clinical sub-populations, where patients with low symptoms severity served as a control group. The conclusions regarding motor activity, facial flatness and congruity should therefore be restricted to the observed clinical population. Second, facial features were extracted from a one-time 15-minute clinical session, and may indicate subjects' temporary state and not their overall condition. Although motor data was collected for a considerable period of 3 weeks on average, it

did not necessarily cover the full course of patients illness, and some drug-induced behaviors may not have been recorded. Finally, our study only focused on the patient's behavior, while taking into account care-giver non-verbal behavior and the interaction between them may offer a great clinical value.

To strengthen and validate our results there is a need for a larger, clinically diverse sample with a drug-monitored population, carried out for a longer period of time. In addition, future analysis should combine different non-verbal features (including vocal expressions, heart rate and skin conductance), comparing between them in order to gain a more holistic representation of patients behavior. The derived measures may than be tested for their relation with cognitive and neural mechanisms of schizophrenia disorder and other psychiatric and neurological conditions.

Before we conclude this discussion, it is important to take a step back considering the possible implications of automatic diagnostic systems in the field of psychiatry. The ability to recognize and response to a diversity of human behaviors was developed over thousands of years in the evolution process. Mental disorders are deficits in the most high level functions of human behavior, and though objective tools are of great need, they may come with a price. Treating psychiatric disorders the same as other medical conditions, while looking for absolute objective measures for diagnosis and treatment, may lessen the broader view of a person as a whole. Both empirical studies and clinical experience emphasize the importance of combining medical treatment with social and clinical therapy. Automatic tools and algorithms are as good as we train them to be, and though they might be very beneficial for quantification and objectivization of certain clinical symptoms, they can not replace a human care giver. Also, manners of privacy and voluntary cooperation must be taken into account. Putting the full weight on such automatic diagnostic tools without taking responsibility for the patient's psychological and clinical condition would be at best unprofessional and at worst malpractice.

Notwithstanding, this dissertation demonstrates the huge potential of using assistive technology for behavioral and affective evaluation. We offer a new way of thinking about facial expressions and motor behavior in schizophrenia, and the concepts introduced here may be of great value in

both clinical settings and future empirical developments. The measures derived in this dissertation may shed light on other neurological and psychiatric conditions, and can be implemented in future diagnostic assisting systems. We believe that technology which puts emphasis on facial dynamics and intensity, while taking into account behavioral and psychological context, may facilitate monitoring of patients, advance detailed evaluation of symptom severity and promote precise adjustments of pharmacological treatment.

ניתוח אוטומטי של התהגות לא מילולית בסכיזופרניה

Automated Analysis of Nonverbal Behavior in Schizophrenia

מאת: טליה טרון מנחה: פרופ' דפנה ויינשל

תקציר

אבחון פסיכיאטרי של הפרעות נפשיות בכלל וסכיזופרניה בפרט מתבסס ברובו על מדדים לא כמותיים ולא אובייקטיביים. התנהגויות לא מילוליות של המטופל כגון תנוחת הגוף, הבעות הפנים ומאפייני הדיבור מהווים חלק אינטגרלי ומשמעותי מתהליך האבחון. עם זאת, המדדים הקיימים להתנהגויות אלו אינם מוגדרים היטב ומרבית האבחון מבוסס על אינטואיציה וניסיון המטפל, דבר המקשה לנטר מטופלים לאורך זמן, להשוות בין מטופלים שונים ולהעריך את השפעות הטיפול. בשנים האחרונות חלה התפתחות משמעותית בניתוח אוטומטי של תנוחות גוף והבעות פנים, וכיום ניתן לחלץ מידע זה באופן אמין יחסית מוידאו ומסנסורים תוך שימוש בתוכנות ייעודיות.

מטרת המחקר המפורט בתזה זו היא לאפיין את ההתנהגות הלא מילולית של חולי סכיזופרניה, ולפתח כלים אוטומטיים לתיאור וניתוח כמותי של התנהגות זו. אנו מתמקדים בהבעות פנים ובפעילות מוטורית, ומשלבים שיטות סטטיסיות קלאסיות עם טכניקות של למידת מכונה (machine learning). באמצעות כלים אלו פיתחנו מגוון רחב של מדדים לא מילוליים, הכוללים את עצמת ההבעה והפעילות המוטורית, הדינמיות שלה, קונסיסטנטיות לאורך זמן והתאמה לסיטואציה. מדדים אלו משמשים אותנו לתאר בצורה מדוייקת יותר את השינויים ההתנהגותיים אצל מטופלים, לבחון את הקשר בין הסימפטומים הלא-מילוליים השונים, ולאפיין כיצד הם משתנים עם שיפור או החמרה במצב הקליני. בנוסף, השתמשנו במדדים הלא-מילוליים על מנת לפתח כלים אוטומטים לזיהוי תת-אוכלוסיות קליניות, להערכה של חומרת סימפטומים ולזיהוי אירועים חריגים אצל מטופלים.

אנו מאמינים כי השיטות והגישות לניתוח התנהגות לא-מילולית המוצגות בתזה זו, יתרמו לתחום המחקר העוסק בחישוביות של הבעה רגשית (affective computing) בכלל, ולתחום האבחון הפסיכיאטרי בפרט. תקוותנו היא לשפר את מהימנות האבחון, לאפשר אפיון וניטור טוב יותר של מטופלים לאורך זמן ולקדם הן את המחקר האמפירי והן את הטיפול הקליני והתרופתי.

עבודה זו נעשתה
בהדרכתה של:
פרופ' דפנה ויינשל

האוניברסיטה העברית בירושלים

ניתוח אוטומטי של התנהגות לא מילולית בסכיזופרניה

מאת
טליה טרון

חיבור לשם קבלת תואר
"דוקטור בפילוסופיה" ב-
מדעי המח : חישוב ועיבוד מידע

הוגש לסנט האוניברסיטה העברית בירושלים
דצמבר 2017 טבת תשע"ח