

# Monitoring the effects of therapeutic interventions in depression through self-assessments

Ines Moragrega,<sup>1</sup> René Bridler,<sup>2</sup> Christine Mohr,<sup>3</sup> Michela Possenti,<sup>4</sup> Deborah Rochat,<sup>3</sup> Judit Sanchez Parramon,<sup>1</sup> Hans H. Stassen<sup>5</sup>

<sup>1</sup>Department of Psychobiology, University of Valencia, Valencia, Spain; <sup>2</sup>Sanatorium Kilchberg, Kilchberg, Switzerland; <sup>3</sup>Department of Psychology, University of Lausanne, Lausanne, Switzerland; <sup>4</sup>Department of Psychology, University of Milano Bicocca, Milano, Italy; <sup>5</sup>Institute for Response-Genetics, Department of Psychiatry, Psychotherapy and Psychosomatics, Psychiatric University Hospital, Zurich, Switzerland

## ABSTRACT

The treatment of major psychiatric disorders is an arduous and thorny path for the patients concerned, characterized by polypharmacy, massive adverse side effects, modest prospects of success, and constantly declining response rates. The more important is the early detection of psychiatric disorders prior to the development of clinically relevant symptoms, so that people can benefit from early interventions. A well-proven approach to monitoring mental health relies on voice analysis. This method has been successfully used with psychiatric patients to 'objectively' document the progress of improvement or the onset of relapse. The studies with psychiatric patients over 2-4 weeks demonstrated that daily voice assessments have a notable therapeutic effect in themselves. Therefore, daily voice assessments appear to be a low-threshold form of therapeutic means that may be realized through self-assessments. To evaluate performance and reliability of this approach, we have carried out a longitudinal study on 82 university students in 3 different countries with daily assessments over 2 weeks. The sample included 41 males (mean age 24.2±3.83 years) and 41 females (mean age 21.6±2.05 years). Unlike other research in the field, this study was not concerned with the classification of individuals in terms of diagnostic categories. The focus lay on the monitoring aspect and the extent to which the effects of therapeutic interventions or of behavioural changes are visible in the results of self-assessment voice analyses.

The test persons showed an over-proportionally good adherence to the daily voice analysis scheme. The accumulated data were of generally high quality: sufficiently high signal levels, a very limited number of movement artifacts, and little to no interfering background noise. The method was sufficiently sensitive to detect: i) habituation effects when test persons became used to the daily procedure; and ii) short-term fluctuations that exceeded pre-specified thresholds and reached significance. Results are directly interpretable and provide information about what is going well, what is going less well, and where there is a need for action. The proposed self-assessment approach was found to be well-suited to serve as a health-monitoring tool for subjects with an elevated vulnerability to psychiatric disorders or to stress-induced mental health problems. Daily voice assessments are in fact a low-threshold form of therapeutic means that can be realized through self-assessments, that requires only little effort, can be carried out in the test person's own home, and has the potential to strengthen resilience and to induce positive behavioural changes.

**Key words:** Psychiatric disorders; depression; psychosomatic disturbances; stress-related health problems; early detection; self-monitoring; biofeedback; university students; voice analysis.

Correspondence: Hans H. Stassen, Psychiatric University Hospital Zurich (KPPP), Institute for Response-Genetics, P.O. Box 1931, CH-8032 Zurich, Switzerland.  
E-mail: k454910@bli.uzh.ch

Citation: Moragrega, I., Bridler, R., Mohr, C., Possenti, M., Rochat, D., Sanchez Parramon, J., & Stassen H. H. (2021). Monitoring the effects of therapeutic interventions in depression through self-assessments. *Research in Psychotherapy: Psychopathology, Process and Outcome*, 24(3), 250-262. doi: 10.4081/ripppo.2021.548

Acknowledgements/funding: this project was funded in part through the 7<sup>th</sup> EU Framework Programme for Research and Technological Development (Grant 248544).

Ethics: the research was conducted in accordance with the Declaration of Helsinki and approved by the local Ethics Committees in Italy (IRB University of Milano-Bicocca), Spain (IRB University of Valencia), and Switzerland (University of Lausanne). All students were informed about the goals of the project and that they can discontinue the 14-day assessment period at any time without giving reasons. All students gave written informed consent.

Received for publication: 10 May 2021.  
Revision received: 16 July 2021.  
Accepted for publication: 7 September 2021.

This work is licensed under a Creative Commons Attribution Non-Commercial 4.0 License (CC BY-NC 4.0).

©Copyright: the Author(s), 2021  
Licensee PAGEPress, Italy  
*Research in Psychotherapy: Psychopathology, Process and Outcome 2021; 24:250-262*  
doi:10.4081/ripppo.2021.548

## Background

### Rationale

Over the past 20 years, stress-related health problems were on the rise worldwide, among which psychosomatic disturbances and impaired mental health play a prominent role. Needless to say, that all forms of impaired mental health contribute to the burden of disability and mortality,

thus having both direct and indirect impacts on length and quality of life. In fact, psychiatric disorders account for 21.2% of years lived with disability worldwide (Mental Health Foundation, 2016). Available treatments, though effective, are incomplete since all treatment options are non-causal, and there is no long-term cure for a considerable proportion of patients: for example, schizophrenic disorders with not less than 50%-60% (e.g., Davidson, 2018), and major depression with 35%-50% (e.g., Pigott, Leventhal, Alter, & Boren, 2010; Safer, 2019).

Today's psychiatric patients are no longer treated with one single medication ('monotherapy') but receive combinations of antidepressants, antipsychotics, mood stabilizers, anxiolytics, antihistamines, and anticholinergics, among others ('polypharmacy'). Psychotherapy without parallel medication is not even considered in the vast majority of cases. At the same time as the polypharmacy approach has become the de facto standard in psychiatry, the percentage of treatment responders has declined dramatically. About 15 years ago, Stassen *et al.* (2007) found in a cross-comparison of 5 antidepressants ( $n=2245$ ) responder rates between 47.5% and 60.9% under monotherapy (Stassen, Angst, Hell, Scharfetter, & Szegedi, 2007) while the responder rates under antipsychotics lay in the range of some 40%.

By contrast, in a recent longitudinal study of 279 inpatients, the patients received an average of  $4.54 \pm 2.68$  concurrent medications, composed of  $3.30 \pm 1.84$  psychotropic drugs, plus  $0.81 \pm 1.13$  drugs to reduce unwanted side effects, plus  $0.43 \pm 0.89$  other somatic drugs. The responder rate was 35% for major depression and 25% for schizophrenic disorders (Stassen *et al.*, 2021), which meant a general drop of about 40% compared to what was the standard 20 years ago. Among the patients suffering from major depression, monotherapy was a rare exception (12.7%), and only a small minority of patients under psychotherapy received psychotherapy alone (2.9%). The figures for outpatients are essentially the same ( $n=363$ ; Löttscher, Angheliescu, Braun, Bridler, & Stassen, 2010).

In tandem with the potentially beneficial effects of psychopharmacological treatments, patients experience significant adverse side effects. In a recent study, 85.7% of patients treated for major depression reported unwanted side effects (31.0% in severe form) and 81.7% of patients treated for schizophrenic disorders (33.1% in severe form) (Stassen *et al.*, 2021). There is no doubt that in many cases the beneficial effects of psychotropic drug treatment do not outweigh the associated risk of adverse side effects. And worst of all, there is no causal treatment of major psychiatric disorders to be expected in the near future.

### Early detection and prevention

Today's treatment of major psychiatric disorders is an arduous and thorny path for the patients concerned, characterized by polypharmacy, massive adverse side effects, modest prospects of success, and constantly declining re-

sponse rates. The more important is the early detection of latent psychiatric disorders prior to the development of clinically relevant symptoms, so that people can benefit from early interventions ('preventing illness instead of treating it') (e.g., Calear, & Christensen, 2010; Albert, & Weibell, 2019). The early detection would not only be to the benefit of sufferers but would also have an enormous socio-economic impact. For example, depression was in 2013 the second leading cause of years lived with disability worldwide (Mental Health Foundation, 2016).

Early detection and prevention necessarily imply the active involvement of people with an elevated risk of developing mental illnesses. Indeed, getting people actively involved and getting them to do something about mental health risks are the most important steps in this context. Here come self-assessments and self-monitoring into play, as self-assessments on a regular basis not only inform the person concerned about changes in health, but also increase health awareness. Specifically, monitoring health through regular self-assessments enables the detection of clinically relevant changes in physical or mental health at early stages ('deterioration', 'improvement'), thus enabling early beneficial interventions (e.g., Calear, & Christensen, 2010; Albert, & Weibell, 2019).

### Self-assessments

A multitude of e-health gadgets on the market allows users to monitor physical health in an 'objective' way through self-assessments of biological data, such as heart rate, heart rate variability, blood pressure (Albus, 2010; Arakawa, 2018; Esler *et al.*, 2008), physical activity (Silanowicka, & Thakuria, 2019), sleep quality, skin conductance, body temperature, and cortisol (Aschbacher *et al.*, 2013; Bhake *et al.*, 2019; Chopra *et al.*, 2019; Lee, Kim, & Choi, 2015; Oswald *et al.*, 2006; Staufenbiel, Penninx, Spijker, Elzinga, & van Rossum, 2013). All this provides educative information about the subjects' lifestyle - did they get enough sleep, did they have enough physical activity, could they count on social contacts, and what have been their bodily reactions in stressful situations.

There are considerably fewer options when it comes to 'objectively' monitoring mental health through self-assessments with regard to psychosomatic disturbances, burn-out conditions, social anxiety, or depressive and schizophrenic disorders. The popular e-health screening instruments on the basis of mental health questionnaires do not really offer a viable solution regarding repeated assessments over a longer period of time. The more so, as these instruments do not yield 'objective' information about health as is the case with biological data.

A well-proven approach to monitoring mental health relies on voice analysis. In fact, human speech production is the result of a joint effort of mind and body. It involves a cascade of steps from utterance planning to final sound production with hundreds of degrees of freedom (Titze, 1994). Attentive listeners can discover a lot about the

physical and mental state of their dialog partners without having to talk about it explicitly during a conversation or on the phone (Braun *et al.*, 2014).

Speaking behaviour such as hectic and abrupt, or delayed and monotonous speech can indicate mental health problems or stress-related adverse reactions, *provided such behaviour persists over a longer time period* (Kraepelin, 1927). The same is true for voice sound characteristics inherent, for example, in a sharp, metallic, or expressionless voice that lacks ‘tonal richness’ in timbre.

Speaking behaviour and voice sound characteristics encompass features like ‘speech flow’, ‘loudness’, ‘intonation’, and ‘vocal timbre’ which can be quantified through acoustic parameters like speaking rate, length of pauses, energy (loudness), vocal pitch, or jitter changes in vocal pitch over a short period of time (Braun *et al.*, 2016; Slavich, Taylor, & Picard, 2019). In the era of machine-learning and artificial intelligence, these measures are often combined into abstract classifiers which inform about presence/absence of certain clinical conditions (Cummins, Baird, & Schuller, 2018; McGinnis *et al.*, 2019).

Typical examples of information inferable directly from a subject’s speaking behaviour and voice sound characteristics, are stress-induced bodily reactions like uneasiness, doziness, weariness from bodily or mental exertion, fatigue, anger, aggression, sadness, grief, and fear, as well as mental distress, and of course psychiatric conditions involving depressive and/or psychotic symptoms (*e.g.*, Johar, 2016).

The voice analysis method has been successfully used in psychiatry for monitoring the time course of recovery among depressive and psychotic patients (Arevian *et al.*, 2020; Faurholt-Jepsen *et al.*, 2016; Hashim, Wilkes, Salomon, Meggs, & France, 2017; Püschel, Stassen, Bomben, Scharfetter, & Hell, 1998; Stassen, Kuny, & Hell, 1998; Stassen *et al.*, 2011; Taguchi *et al.*, 2018; Wang *et al.*, 2019). In this approach, each patient’s vocal baseline at entry into the observation period serves as a reference line from which deviations in positive and negative direction are measured. In fact, as speech production can be influenced by numerous endogenous and exogenous factors, the resulting speech exhibits ‘natural’ variation and sometimes ‘significant’ deviations from what is ‘normal’ (Garrett, & Healey, 1987). This natural variation fluctuates around a fictitious ‘baseline’ (resting position), which in turn can take on different ‘levels’ depending on the physical and mental state of the speaker (or simply in the course of the day due to diurnal rhythms).

When subdivided in shorter epochs, a two-minute speech recording provides enough information to calculate an estimate of the current baseline for each parameter used to quantify speaking behaviour and voice sound characteristics, for example, by regression over epochs. Of specific interest is the development of a subject’s baselines over a longer time period (several days or weeks). This model has been successfully tested, for example, on depressive patients over 4 weeks with voice assessments

at 2-day intervals along with psychopathology ratings immediately after voice assessments (HamD-17). For patients showing response to treatment, single case analyses revealed a close correlation of  $r \geq 0.8$  between the patients’ psychopathology scores and voice sound characteristics during recovery (Stassen *et al.*, 2011; Stassen *et al.*, 2007; Stassen, Kuny, & Hell, 1998). The study also showed that the method is particularly useful for the early detection of relapses in patients who have successfully overcome their depression.

Unlike other studies in the literature, we are not at all concerned with classifying people in terms of mental illnesses in the sense of a clinical diagnosis. In fact, attempts to derive a clinical diagnosis through just a few voice recordings - as sometimes proposed in the literature - can be grossly misleading with false-positive and false-negative classification errors exceeding 20% by far. By contrast, our focus lies on the monitoring aspect, which has proven very useful in longitudinal studies with psychiatric patients where the time course of the improvement was recorded in an ‘objective’ way by a technician.

We think that this monitoring aspect with all its benefits can also be realized with a self-assessment approach. The more so, as we learned from our studies with psychiatric patients that routinely performed voice analyses have a therapeutic value by themselves. The vast majority of patients really liked to participate and successively developed a curiosity for possible positive changes in their state of health and course of illness. Accordingly, we expect that a larger proportion of outpatients under therapy, or of subjects with an elevated risk of developing depressive disorders, get interested in the extent to which the effects of therapeutic interventions or of behavioural changes become visible through the results of self-assessment voice analyses. Indeed, getting involved is the most important step towards the early detection and prevention of mental health issues.

### Longitudinal self-assessment study

To evaluate the performance of the voice analysis method in *self-assessments*, we designed a longitudinal study with daily assessments over 14 days. Results of self-assessment voice analyses are presented to the user in form of directly interpretable biological quantities, that is, in the form of immediate biofeedback. Focus was laid on university freshman students in the age range of 17-27 years (Stallman & Shochet, 2009; Stallman, 2010): i) Freshman students encounter significant levels of chronic stress<sup>1</sup> over quite a long time caused by competition in classroom, tight schedules, frequent exams, and adaptation to new environments (Corley, 2013; Hunt & Eisenberg, 2010; Li, Lindsey, Yin, & Chen, 2012); ii) Some

<sup>1</sup> Under ‘stress’ we understand the state of bodily and/or mental strain which is caused by environmental or endogenous conditions that have measurable impacts on both body and mind in the general population.

12-18% of subjects of the general population show insufficient coping behaviour under chronic stress, thus possessing an elevated risk of developing mental health problems (Delfino *et al.*, 2015; Mohr *et al.*, 2014; Zhang *et al.*, 2019); iii) 75% of subjects with major psychiatric disorders have their onset in this age range (Kessler *et al.*, 2007); iv) Freshman students are a central target population for self-monitoring approaches, given the number of students who seek psychological counselling services on university campuses and their above-average willingness to cooperate in health issues.

### Aims and hypotheses

Our study aimed at the following topics: i) overall compliance with the voice analysis method encompassing daily voice assessments; ii) potential socio-cultural differences in compliance; iii) data quality of self-assessments; iv) inter-relation between basic coping behaviour and general health; v) stability of speaking behaviour and voice sound characteristics across daily self-assessments over 14 days<sup>2</sup>; vi) interpretation of longitudinal results; and vii) sensitivity of self-assessments regarding the detection of deviations from ‘normality’.

## Materials and methods

### Sample composition

Our sample was comprised of 83 students (42 males, 41 females), recruited at three culturally different sites with three ‘syllable-timed’ languages: the Universities of Lausanne/Switzerland (*French*), Milano/Italy (*Italian*), and Valencia/Spain (*Spanish*). All students were informed about the goals of the project and that they can discontinue the 14-day assessment period at any time without giving reasons. They were then invited to fill out the two self-report questionnaires that evaluate basic coping behaviour and general health: i) the 28-item Coping Strategies Inventory (COPE); and ii) the 63-item Zurich Health Questionnaire (ZHQ)<sup>3</sup>. Subsequently, the students were asked to carry out daily voice assessments over 14 days using two short pieces of speech: ‘counting out loud’ and ‘reading out loud standard text’ (standardized 2-minute self-assessments). A minimum of 12 self-assessments was requested. Of the 83 students in the study, a total of 80 met this condition (96.4%), of which the majority completed exactly 12 assessments as requested (not necessarily on consecutive days), while some had 1-3 additional assessments.

<sup>2</sup> The distinct inter-individual differences of human voices along with the stability of speaking behaviour and voice sound characteristics over time enable the recognition of persons through their voices.

<sup>3</sup> These instruments are available in 6 different languages (*English, French, German, Italian, Spanish, and Chinese*) and can be downloaded for free from website: <https://ifma-health.com/instruments.php>

## Measures and procedures of measurement

### Self-report questionnaires

For the analysis of the COPE and ZHQ questionnaires, we relied on the results of previous normative studies on 2,517 students, where structural analyses had shown that the information assessed through the COPE instrument can be summarized by two scales reflecting socio-culturally independent personality traits ‘Activity-Passivity’ and ‘Defeatism-Resilience’, while the information assessed through the ZHQ instrument can be summarized by eight scales: i) tobacco consumption; ii) alcohol consumption; iii) regular use of medicine; iv) illegal drugs; v) regular exercises; vi) impaired physical health; vii) psychosomatic disturbances; and viii) impaired mental health.

### Voice analysis approach

Of central importance for the methodological approach we used for voice analysis was the quantification of speaking behaviour and voice sound characteristics by means of directly interpretable variables, which were provided to the test persons in the sense of immediate biofeedback. This is in fact readily possible: Even though rhythm, stress, and intonation (‘prosody’) greatly influence the *content* of speech, speaking behaviour and voice sound characteristics can be comprehensively described by a few major features: i) speaking behaviour in terms of ‘speech flow’, ‘loudness’, and ‘intonation’, while ii) voice sound characteristics relate to the distribution and intensity of ‘overtones’ that make up a speaker’s individual vocal ‘timbre’ (*e.g.*, bright *versus* dark colour).

Speech flow describes the speed at which utterances are produced, along with the number and duration of temporary breaks in speaking. Loudness reflects the amount of energy associated with the articulation of utterances and, when regarded as a time-varying quantity, the speaker’s dynamic expressiveness. Intonation is the manner of producing utterances with respect to rise and fall in pitch and leads to tonal shifts in either direction of the speaker’s mean vocal pitch. Overtones are the higher tones which faintly accompany a fundamental tone, thus being responsible for the tonal diversity of sounds. Stress, psychosomatic disturbances and psychiatric disorders influence a person’s overtone pattern in characteristic ways.

One observes short-term fluctuations due to the test person’s interactions with the immediate environment and long-term changes of a few days or even weeks caused by mental health problems or adverse reactions to chronic stress. The line between ‘natural fluctuations’ and ‘significant changes’ essentially depends on spoken language, gender, age, as well as type of speech (*e.g.*, Cohen, Renshaw, Mitchell, & Kim, 2016; Lortie, Thibeault, Guitton, & Tremblay, 2015), such as automatic speech (‘counting’), emotionally neutral, or emotionally charged speech. In order to decide on the significance of changes, we re-

lied on normative data from repeated assessments on 613 healthy subjects stratified according to: i) gender and age: age range 18-65 years; ii) the 'stress-timed' languages *English* and *German*; and iii) the 'syllable-timed' languages *French*, *Italian*, and *Spanish* (Braun *et al.*, 2014). These normative data from healthy subjects were complemented by threshold values from psychiatric patients under treatment (n=598: 350 males, 248 females), thus providing a quite reliable database for our AI procedures.

The above approach yields, in the sense of 'biofeedback', directly interpretable quantitative voice parameters. From these parameters, a test person can directly deduce what is going well, what is going less well and where there is a need for further improvement. Thus, a test person can iteratively improve mental health problem areas through active behavioural changes, for example, how to calm down, to gain distance, to reduce tension, to concentrate on one single task for several minutes, to have the energy to pursue a goal, to communicate better with peers, to work on a project for a longer period of time and to complete it successfully. This in contrast to methods that yield more abstract classifiers regarding a test person's mental health status (cf. Cummins *et al.*, 2015; Grabowski *et al.*, 2019).

We recorded speech signals as WAV-files with a sampling rate of 48 kHz at a 16-bit resolution. The resulting time series were automatically subdivided into utterances and pauses by means of an AI-optimized, language-specific segmentation algorithm. We then used Discrete Fourier Transformations (DFTs) to calculate 'spectra' from 'pure' utterances with pauses being skipped. Spectral analyses relied on a tonal approach with a quartertone resolution over 7 octaves in the frequency range of 64-8192Hz, yielding spectra of 168 equally spaced quartertones. The tonal approach was chosen because pitch (perceptual quantity) depends logarithmically on frequency (physical quantity). Finally, the intermediate data were used to extract parameters that quantify speaking behaviour and voice sound characteristics.

All the above mentioned steps are implemented in a platform-independent JAVA application that runs on any system that supports JAVA, as well as on ANDROID devices. The WAV-files stored on self-assessment devices can be uploaded via internet to the large-scale server of our research group in order to perform detailed analyses on the population level.

### Statistical analyses

For our statistical analyses, we relied on the Statistical Analysis Software SAS/STAT 9.4 by SAS Institute Inc., Cary NC (USA). The evaluation of the self-report questionnaires COPE and ZHQ, as well as the analysis of the voice recordings are standardized and encapsulated in SAS macros that were developed within the scope of our previous studies for quality control purposes. For AI applications and Neural Networks, we relied on SPSS 25

(Neural Networks) by IBM Software, Armonk, NY (USA), SAS Enterprise Miner 15.1 (PROC HP-NEURAL), in combination with a proprietary program (NNA) developed by our research group.

## Results

### Topics 1, 2: Compliance

Of the 83 students recruited n=27 were from the French speaking part of Switzerland [14 males, 13 females; mean age: 24.2±3.2 years], n=26 from Italy [13 males, 13 females; mean age: 23.5±3.5 years], and n=30 from Spain [15 males, 15 females; mean age: 21.3±1.8 years]. Totally 82 students (98.8%) successfully completed the envisaged study period, despite the fact that they could discontinue the 14-day assessment period at any time without giving reasons. There was only one premature withdrawal. Of the expected 82·12=984 self-assessments, 22 (2.2%) were missing for diverse reasons. Hence, the three populations under investigation showed an over-proportionally good adherence to the study protocol and, in particular, to the daily voice analysis scheme. No socio-cultural differences were found in this respect.

### Topic 3: Data quality in self-assessment voice analyses

Due to missing data, our database comprised only 962 data records (missing data rate of 2.2%). In a first step, we calculated the test persons' individual speaking behaviour and voice sound characteristics separately for each assessment. In the second step, the repeated assessments were combined into a longitudinal model and evaluated for each individual test person. A minimum of five assessments were required for the longitudinal model. All 82 test persons met these requirements, thus underlining again the test persons' good compliance with the study protocol. The voice recordings were of generally high quality: sufficiently high signal levels, a very limited number of movement artifacts, and little to no interfering background noise. However, due to excessive distortions of unknown origin, we had to exclude 17 recordings (1.8%) from further analyses.

### Topic 4: Self-assessment of basic coping behaviour and general health

The same over-proportionally good compliance was also evident in the analysis of the internal consistency of the students' self-report questionnaires by means of the imbedded control items. No consistency violation showed up in the empirical data, thus suggesting that the COPE and ZHQ questionnaires were completed in a very cooperative way. The scales computed from the questionnaire data were compatible with those of the normative studies and displayed a similar between-subject variation. Descriptive statistics derived from the combined data of the

COPE and ZHQ questionnaires yielded typical male-female differences regarding consumption behaviour: male students consumed more tobacco, alcohol, and illegal drugs, whereas female students showed higher regular use of medicine (Table 1).

Unexpectedly, the female students were found to exhibit a more ‘active’ coping behaviour compared to their male fellow students. No gender differences showed up for the Body Mass Index (BMI). As to the ZHQ scales, we found the typical correlations between ‘impaired mental health’ on the one hand, and i) ‘psychosomatic disturbances’ ( $r=0.330$ ;  $P=0.0023$ ); ii) ‘regular use of medicine’ ( $r=0.239$ ;  $P=0.0239$ ); and iii) ‘regular exercises’ ( $r= -0.229$ ;  $P=0.0376$ ) on the other.

Regarding the inter-relation between basic coping behaviour and general health, we actually found the expected correlation between ‘defeatism’ and ‘psychosomatic disturbances’ ( $r=0.292$ ;  $P=0.0075$ ), yet the correlation with ‘impaired mental health’ did not reach significance ( $r=0.171$ ;  $P=0.1224$ ). Detailed analyses indicated that an over-proportionally large number of students had ‘resilience’ scores well above the average, along with an unexpectedly low rate of subjects with insufficient coping behaviour: 8.4% rather than 15-18% as suggested by the epidemiologic data (Figure 1).

#### Topic 5: Stability of speaking behaviour and voice sound characteristics over time

Distinct between-subject differences along with a remarkable stability of voice patterns over time enable the recognition of persons through their voices. We used this property as independent verification of the data quality of

self-assessments. Already a simple combination of 12 speech parameters yielded a recognition rate of 90.2% of uniquely identified subjects in our sample (74 out of 82 test persons). A finding, which again underlined the remarkably good data quality observed with the self-assessments.

Among 70 test persons (85.4%) the within-subject fluctuations of speech parameters over 14 days were found to be compatible with those observed in normative studies with repeated assessments at 14-day intervals ( $n=613$ ). By contrast, 12 test persons (14.6%) showed significantly higher within-subject fluctuations of speech parameters

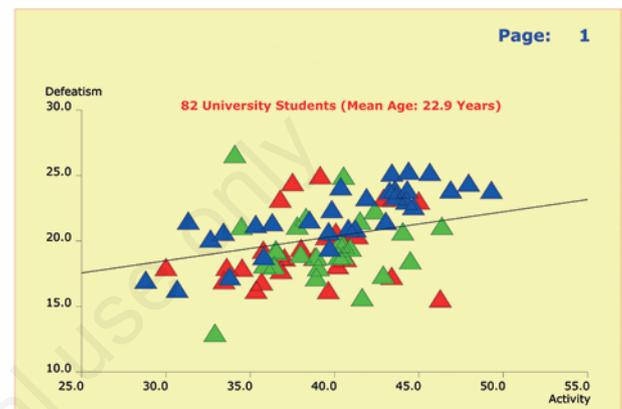


Figure 1. Scatter plots of the raw scores ‘activity’ (x-axis) versus ‘defeatism’ (y-axis) as derived from the COPE data of 82 university students from the French speaking part of Switzerland (red triangles), Italy (green triangles), and Spain (blue triangles). A sufficiently large between-subject variation is a necessary prerequisite to study the interrelations with health factors at sufficiently high resolution.

**Table 1. Male-female comparisons revealed the ‘usual’ differences regarding consumption behaviour: males consumed more tobacco, alcohol, and illegal drugs, whereas females showed a higher regular use of medicine. Yet unexpectedly, the female students were found to exhibit a more ‘active’ coping behaviour compared to their male fellow students, where ‘active’ is best described through COPE items like ‘turning to work’, ‘getting help and advice from other people’, or ‘coming up with a strategy’. No differences showed up for the body mass index (non-significant results with  $P \leq 0.05$  were labeled ‘n.s.’).**

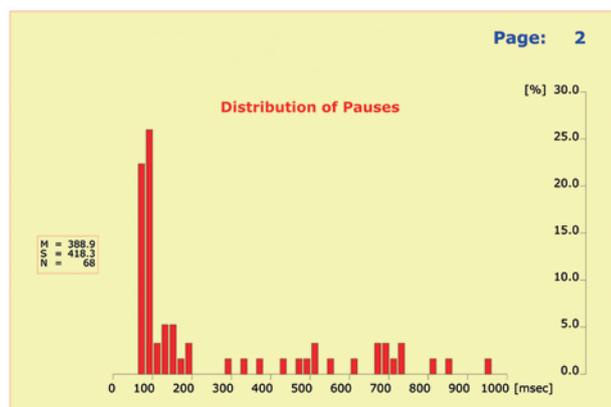
Male-female differences			
	Males	Females	Significance
Tobacco consumption	2.698±4.1	1.057±2.3	$P=0.0277$
Alcohol consumption	6.270±4.2	4.228±4.1	$P=0.0283$
Regular use of medicine	1.726±2.9	2.744±2.7	n.s.
Illegal drugs	3.254±4.5	0.244±0.9	$P<0.0001$
Impaired physical health	0.801±0.9	0.732±0.9	n.s.
Psychosomatic disturbances	3.108±1.9	3.902±2.4	n.s.
Impaired mental health	6.852±2.8	7.290±3.3	n.s.
Regular exercises	8.730±3.8	9.187±4.0	n.s.
Cope activity	38.6±4.9	40.8±3.4	$P=0.0225$
Cope defeatism	20.7±2.7	20.8±3.0	n.s.
Body mass index	22.7± 3.6	21.4±3.5	$P=0.1047$

over the 14-day study period than expected by chance. These over-proportionally large fluctuations may either be health-related or caused by the test persons' momentary emotional state, by stress, lack of concentration, sleepiness, fatigue, or centrally acting substances like alcohol.

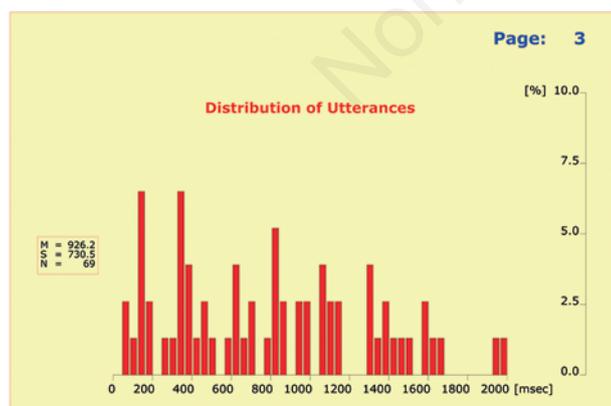
### Topic 6: Interpretation of longitudinal voice analysis results

For each individual test person, the longitudinal analysis yielded the following directly interpretable speech and voice sound characteristics in the form of 'biofeedback':

i) *The variation of pause duration*: a healthy relaxed



**Figure 2. Variation of Pause Duration:** a relaxed speaker presenting a text produces a large variety of pauses of different length and separates some text sections by longer pauses. This in contrast to speakers with mental health problems, in particular patients with depressive symptoms, who tend to present a text in a more monotone, automatized way that lacks variation ('M' means mean pause duration, 'S' standard deviation, and 'n' the number of pauses in the spoken text).



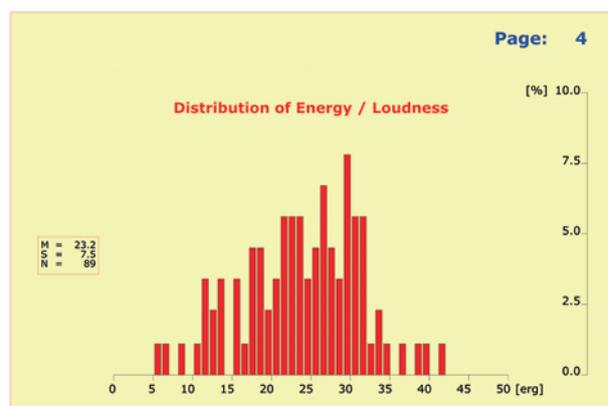
**Figure 3. Variation of Utterance Duration:** when presenting a text, relaxed speakers typically vary the speed by which they produce utterances to make their speech more attractive and interesting. Speakers with mental health problems and speakers under chronic stress wouldn't do this ('M' means mean utterance duration, 'S' standard deviation, and 'n' the number of utterances in the spoken text).

speaker presenting a text produces a large variety of pauses of different length, and typically separates main text sections ('phrases') by longer pauses (Figure 2). This in contrast to speakers with mental health problems, in particular patients with depressive symptoms, or speakers under chronic stress who tend to present a text in a more monotone, automatized way that lacks variation;

ii) *The variation of utterance duration*: when presenting a text, healthy relaxed speakers typically vary the speed by which they produce utterances in order to make their speech more attractive and interesting (Figure 3). Speakers with mental health problems and speakers under chronic stress wouldn't do this;

iii) *The variation of loudness (energy)*: when presenting a text, healthy relaxed speakers typically vary loudness (dynamic expressiveness) in order to make their speech more attractive and interesting. The width of the distribution of loudness (energy) reflects the speaker's dynamic expressiveness (Figure 4). In contrast to healthy subjects, patients suffering from affective disorders, in particular depression, speak in a low voice, slowly, hesitatingly, monotonously, sometimes stuttering or whispering. During recovery, however, patients regain their energy and dynamic expressiveness (Stassen, Kuny, & Hell, 1998; Stassen *et al.*, 2011);

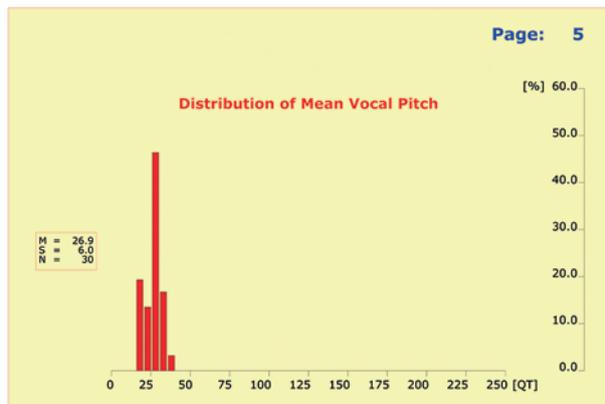
iv) *The variation of vocal pitch (intonation)*: intonation is the manner of producing utterances with respect to rise and fall in pitch, leading to tonal shifts in either direction of the speaker's mean vocal pitch. The 'broader' the variation around the speaker's mean vocal pitch the 'richer' the intonation. The almost complete lack of intentionally used intonation, coupled with irregular jumps in vocal pitch, is typical of the impaired speech in patients with acute schizophrenic symptoms (Lott,



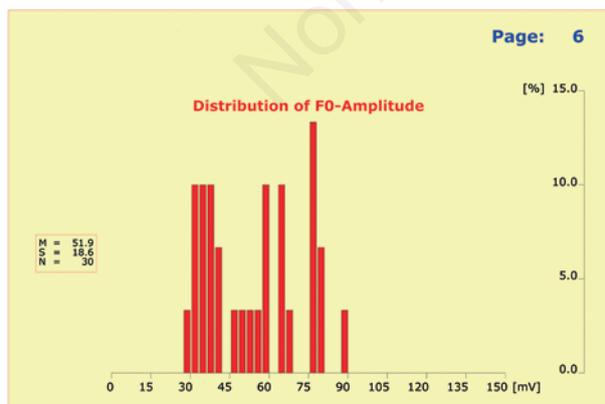
**Figure 4. Variation of Loudness (Energy):** when presenting a text, relaxed speakers typically vary loudness (dynamic expressiveness) to make their speech more attractive and interesting. The width of the distribution of loudness (energy) reflects the speaker's dynamic expressiveness ('M' means mean energy per second, 'S' standard deviation, and 'n' the number of seconds used for the spoken text).

Guggenbühl, Schneeberger, Pulver, & Stassen, 2002). In this context, however, it is worth noting that intonation is qualitatively quite different from a ‘trembling’ voice caused, for example, by shaking involuntarily from cold, by nervousness, fear, or excitement (jitter). Figure 5 shows a speaker with a fairly good intonation, where there is some room for improvement (monitored through self-assessments);

- v) *The variation of F0-amplitude*: this variation is an indicator of the ‘richness’ of a speaker’s vocal timbre.



**Figure 5. Variation of Vocal Pitch (Intonation):** intonation is the manner of producing utterances with respect to rise and fall in pitch. It leads to tonal shifts in either direction of the speaker’s mean vocal pitch. The ‘broader’ the variation around the speaker’s mean vocal pitch the ‘richer’ the intonation. The example shows a typical speaker with fairly good intonation, where there is some room for improvement (‘M’ means mean vocal pitch in quarter tones for the 4 octaves [55-110Hz], [110-220Hz][220-440Hz][440-880Hz], i.e. 48 quarter tones per octave, ‘S’ standard deviation, and ‘n’ the number of segments underlying pitch estimation).



**Figure 6. Variation of F0-Amplitude:** this variation is an indicator of the ‘richness’ of a speaker’s voice sound. A narrow distribution typically means deficiency of emotions and empathetic feelings. By contrast, a broad distribution suggests a lively and mindful person (‘M’ means mean F0-Amplitude, ‘S’ standard deviation, and ‘n’ the number of segments underlying F0 estimation).

Subjects under chronic stress tend to have a ‘sharp’, sometimes ‘metallic’ voice sound, yet regain their bright and full timbre when relaxing. Among affectively disturbed subjects, particularly in depression, a narrow distribution typically means deficiency of emotions and empathetic feelings. A broad distribution, by contrast, suggests a lively and mindful person (Figure 6);

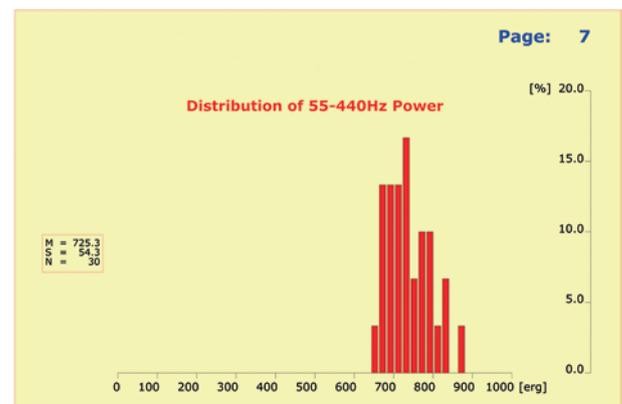
- vi) *The variation of 55-440 Hz power*: this variation is another measure of the tonal ‘richness’ of a speaker’s voice (Figure 7). Reduced variation indicates a more monotonous production of utterances caused, for example, by a lack of energy or concentration, being away with the fairies, sorrowfulness, weariness, or fatigue.

When interpreting the results of voice analyses, it is important to understand that significant deviations from ‘normality’ in a single voice assessment are not a sufficient indication for any elevated risk. In fact, short-term fluctuations in mood are constituents of human life, reflecting: i) interactions with the environment; and ii) usually appropriate reactions to the daily grind.

*If and only if significant deviations from normality show cumulative occurrence, or persist over a longer time period, then the person concerned should be on the alert and it might become necessary for him/her to do something about, if necessary, to call on professional help.*

#### Topic 7: Sensitivity of self-assessment voice analyses

In the single case analyses, the vast majority of students displayed stable speech parameters over the entire study period even though short-time fluctuations occasionally reached significance. An example is shown in Figure 8 where the test person displayed a slight reduction in loudness between days 10 and 13 with otherwise almost constant values. Overall, we found little to no evidence of stress-related psychosomatic disturbances or acute psy-



**Figure 7. Variation of 55-440 Hz Power:** this variation is another measure of the tonal ‘richness’ of a speaker’s voice. Reduced variation indicates a more monotonous production of utterances caused, for example, by sorrowfulness, weariness, or fatigue (‘M’ means mean energy in the frequency range 55-440 Hz, ‘S’ standard deviation, and ‘n’ the number of segments underlying 55-440 Hz Power estimation).

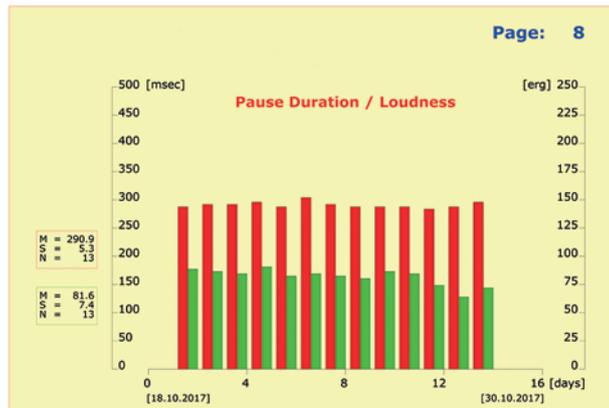
chiatric conditions among the students throughout the 2-week study period.

None of the students exhibited significant changes in speaking behaviour and voice sound characteristics that persisted over a longer time period as one observes with psychiatric patients during recovery. When we observed significant deviations from ‘normality’ then these deviations were short-lived, likely caused by factors like sleepiness, lack of concentration, being away with the fairies, sorrowfulness, or weariness. An example is given in Figure 9.

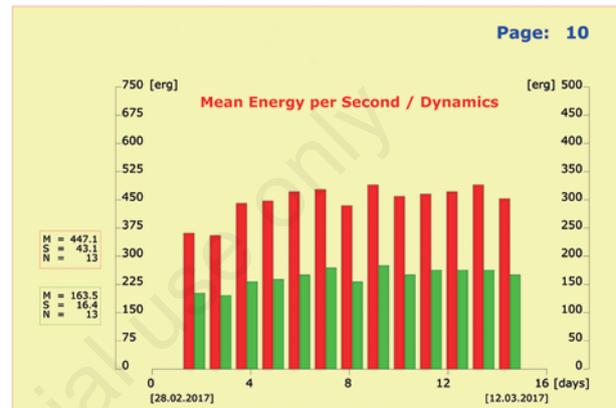
Expectedly, we found habituation effects in a subgroup of test persons. Specifically, some students showed

systematic trends towards shorter pauses and greater loudness when they got more and more used to the test procedure. The same was true for dynamic expressiveness and intonation which steadily improved in a number of test persons once they became familiar with the experimental situation (Figures 10 and 11).

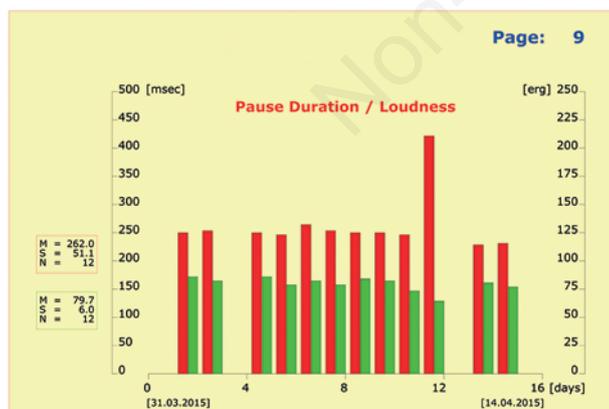
The sensitivity of the voice analysis method regarding the detection of mental health issues or stress-related behaviour is shown in Figure 12, where the test person exhibited some short-lived reductions in intonation during the observation period, reaching significance on day 6. Such reductions suggest physical or psychological reac-



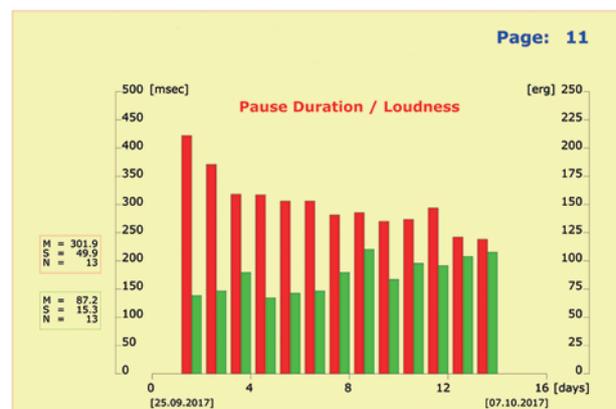
**Figure 8. Pause Duration (red bars)/Loudness (green bars):** Despite their great stability over time, the speech parameters ‘Pause Duration’ and ‘Loudness’ often show a systematic trend toward shorter pauses and greater loudness when speakers get used to the test (‘M’ means mean value of pause duration and loudness, respectively, ‘S’ standard deviation, and ‘n’ the number of repeated assessments [days]).



**Figure 10. Energy (red bars) and dynamics (green bars) over an observation period of 14 days:** Speaking behaviour shows a systematic trend towards higher energy values and greater dynamic variation as a function of time. This finding is likely due to habituation effects (‘M’ means mean value of energy and dynamics, respectively, ‘S’ standard deviation, and ‘n’ the number of repeated assessments [days]).



**Figure 9. Pause Duration/Loudness:** Pause duration (red bars) and loudness (green bars) over an observation period of 14 days: Speaking behaviour is virtually unchanged over time except for day 12 with longer pauses and a lower voice (subject may have been tired). No self-assessments on days 4, 13 (‘M’ means mean value of pause duration and loudness, respectively, ‘S’ standard deviation, and ‘n’ the number of repeated assessments [days]).



**Figure 11. Pause Duration/Loudness:** Pause duration (red bars) and loudness (green bars) over an observation period of 14 days: Speaking behaviour shows a systematic trend towards shorter pauses and greater loudness as a function of time. This finding may indicate habituation effects (‘M’ means mean value of pause duration and loudness, respectively, ‘S’ standard deviation, and ‘n’ the number of repeated assessments [days]).

tions to some life events at that time. In addition to the reduced intonation on day 6, the test person's mean vocal pitch showed a tonal shift towards lower frequencies as well, which was probably due to fatigue (Figure 12).

In summary, by means of directly interpretable results derived from self-assessment voice analyses it was readily possible: i) to detect habituation effects; and ii) to detect short-term fluctuations that exceeded pre-specified age-, gender-, and language-specific thresholds. This should make it equally possible to use regularly repeated self-assessments to detect a longer-lasting deterioration in the mental health at an early stage, for example, among patients who are at risk of relapsing from a depression that has just been overcome, or among persons with an elevated risk of stress-related health problems. Self-assessments can, of course, also be used to monitor steady improvements in mental-health during recovery, as one has seen in the studies with hospitalized psychiatric patients over 2-4 weeks (where voice assessments, however, were carried out in a speech lab by technicians).

## Discussion

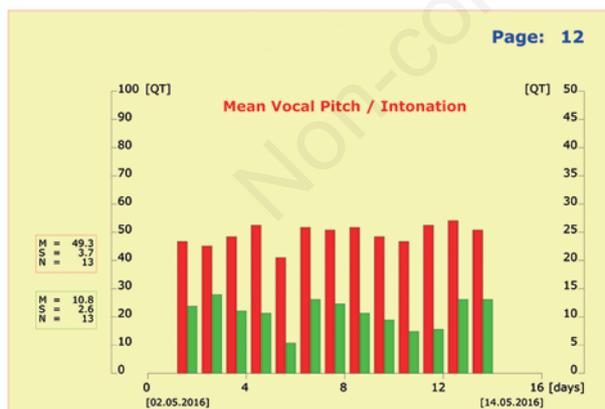
The current standard of treatment in psychiatry is based on polypharmacy and leads primarily to a wide range of undesirable side effects, without providing greater therapeutic benefits to patients. This unsatisfactory situation is due to the fact that there is no causal therapy in psychiatry, while available antidepressants and antipsychotics that differ greatly in their biochemical design and primary site of pharmacological action display virtually

the same insufficient efficacy (Stassen *et al.*, 2007). Since there will be no causal treatment in the near future, we think it is time for psychiatry to rethink its treatment strategies, which are far too one-sidedly fixated on psychopharmacology and also pay far too little attention to the early detection of signs and symptoms in subjects with an elevated risk of relapse, despite the many patients with recurrent episodes. In line with this, we advocate that more attention be paid to the early detection of longer persisting deviations from 'normality' in subjects with an elevated vulnerability to psychiatric disorders or to stress-induced health problems. In this way, many unnecessary treatments and hospitalizations could be avoided.

Inevitably, the success of early detection and prevention procedures depends on the cooperation of the subjects concerned. Accordingly, we have proposed an easy-to-use and well-proven procedure with daily 2-3 minute voice recordings which has been successfully used with psychiatric patients when 'objectively' documenting the progress of improvement or the onset of relapse. From these latter studies over 2-4 weeks one can learn that daily voice assessments have a notable therapeutic effect in themselves (Braun *et al.*, 2016; Stassen *et al.*, 2011). Therefore, daily voice assessments can be regarded as a low-threshold form of therapeutic means that can be realized through self-assessments, requires only little effort, can be carried out in the test person's own home, and has the potential to strengthen resilience and to induce positive behavioural changes.

In this study, we tested performance and reliability of the self-assessment voice analysis method in a home environment setting. Our results surpassed expectations by far. Firstly, the study demonstrated the high, socio-culturally independent compliance of test persons with the daily 2-3 minute self-assessments. Secondly, the quality of the data aggregated by self-assessments was generally high in terms of signal levels, movement artifacts, and background noise. Thirdly, self-assessments were capable to measure the distinct stability of speaking behaviour and voice sound characteristics over time as the vast majority of test persons (85.4%) displayed within-subject fluctuations over time compatible to those found in normative studies carried out in a speech lab. Finally, self-assessments were sufficiently sensitive: i) to detect habituation effects when test persons gained routine and became more familiar with the daily procedure; and ii) to pick up short-term fluctuations that exceeded pre-specified normative cut-off values and reached significance.

In summary, the proposed self-assessment approach was found to be well-suited to serve as a health-monitoring tool for subjects with an elevated vulnerability to psychiatric disorders or to stress-induced mental health problems, as well as for outpatients under therapy. Preferably in combination with one of the popular fitness trackers that continuously record physical activity, heart rate and sleep quality (Schuch *et al.*, 2018).



**Figure 12. Vocal Pitch (red bars)/Intonation (green bars):** the test person exhibits short-term reductions in intonation during the observation period (reaching significance on day 6), thus suggesting physical or psychological reactions to some life events at that time. Additionally, the test person's mean vocal pitch showed a tonal shift towards lower frequencies as well, which was probably due to fatigue ('M' means mean vocal pitch and intonation, respectively, in quarter tones for the 4 octaves [55-110Hz], [110-220Hz][220-440Hz][440-880Hz], i.e. 48 quarter tones per octave, 'S' standard deviation, and 'n' the number of repeated assessments [days]).

Given the exceptionally good compliance seen in this study along with the sensitivity of the self-assessment method to longitudinal changes in speaking behaviour and voice sound characteristics, we expect that a larger proportion of subjects under risk will participate in such programs. In particular, we expect that these subjects get interested in the extent to which the effects of therapeutic interventions or of behavioural changes are visible in the results of voice analyses. Indeed, behavioural changes such as becoming calmer and more relaxed, improved stress management, regular physical activity, and getting enough sleep can be considered as effective therapeutic interventions, the effect of which can be made visible through self-assessment voice analyses as demonstrated by this study.

The primary target population is college/university freshman students, given the steadily increasing number of students who seek psychological counselling services and drop out prematurely for psychosomatic or psychiatric reasons without graduating. Several studies have already carried out with totally 3142 freshmen students at six socio-culturally different sites in the U.S., Europe, South America, and China (Delfino *et al.*, 2015; Mohr *et al.*, 2014; Zhang *et al.*, 2019) in order to identify students with an elevated risk of developing stress-related health problems. These studies clearly benefitted from the students' above-average willingness to cooperate in health issues.

Another promising target group is that fast growing part of the general population for which the 'Quantified Self' is inherent part of daily life ('lifelogging'). The self-assessment voice analysis tool fits perfectly in the lifelogging world. From the health policy perspective this is of major interest as it goes hand in hand with other public attempts to promote health ('resilience') rather than to cure illness. By contrast, it is probably much more difficult to find physicians and psychiatrists in private practice who make patients with latent psychiatric or psychosomatic disorders aware of the benefits of self-monitoring, and to convince them that active involvement is the most important step in successfully dealing with this kind of health problems.

To promote the idea of early detection and prevention, screening tools are available on the Internet free of charge and without advertising, so that anyone can check whether he/she has an elevated vulnerability to health problems under chronic stress ('<https://ifma-health.com/>') [available languages: *English, French, German, Italian, Spanish, Chinese*]. Similarly, a self-assessment voice app is available on the Internet free of charge and without advertising ('<https://play.google.com/store/apps/details?id=ch.uzh.ifrg.voxapp>') [available languages: *English, French, German, Italian, and Spanish*]. Indeed, the socio-economic impact of health promotion can be enormous.

## Conclusions

Impaired mental health is a heavy burden for the affected persons, their families and the working environ-

ment. The treatment of such disorders is an arduous and thorny path for sufferers, characterized by polypharmacy, massive adverse side effects, modest prospects of success, and constantly declining response rates. The more important is the early detection of latent psychosomatic or mental health problems prior to the development of clinically relevant symptoms.

The involvement of the persons concerned is of central importance in this context and includes self-monitoring and self-assessments as vital tools. Our study demonstrated the efficacy of the voice analysis tool in self-assessments, thus clearing the way for its use as a self-monitoring tool by subjects with an elevated vulnerability to psychiatric disorders, psychosomatic disturbances, or stress-induced adverse reactions. The tool yields, in the sense of 'biofeedback', directly interpretable quantitative parameters from which the test person can directly deduce what is going well, what is going less well, and where there is a need for improvement. With this type of biofeedback, the test person can iteratively improve problem areas through active behavioural changes, for example, to calm down, to gain distance, to reduce tension, to concentrate on one single task for several minutes, to have the energy to pursue a goal, or to complete a project successfully. Clearly, this kind of active involvement can be very helpful in dealing successfully with latent psychosomatic or mental health risks.

## Limitations and future directions

There are obvious limitations of this study: i) the results were derived from 'syllable-timed' languages only (*French, Italian and Spanish*) where every syllable within phrases takes roughly the same amount of time; the analysis of 'stress-timed' languages, such as *English and German*, remains to be done; ii) all test persons were students in the age range of 17-24 years and familiar with the use of all kinds of e-technology. Other population groups might show a less favourable compliance with daily self-assessment over a longer time period. However, we are confident that the proposed self-assessment approach works equally well for people over 24 years of age.

As a matter of course, we continue extending our database so that we will be able to verify the method's performance for additional population groups in the near future. The planned extensions will in particular also include studies involving the 'stress-timed' languages *English and German*.

## References

- Albert, N., & Weibell, M. A. (2019). The outcome of early intervention in first episode psychosis. *International Reviews of Psychiatry*, 31(5-6), 413-424. doi:10.1080/09540261.2019.1643703.
- Albus, C. (2010). Psychological and social factors in coronary heart disease. *Annals of Medicine*, 42(7), 487-494. doi:10.3109/07853890.2010.515605.

- Arakawa, T. (2018). Recent research and developing trends of wearable sensors for detecting blood pressure. *Sensors (Basel)*, *18*(9), 2772. doi:10.3390/s18092772.
- Arevian, A. C., Bone, D., Malandrakis, N., Martinez, V. R., Wells, K. B., Miklowitz, D. J., & Narayanan, S. (2020). Clinical state tracking in serious mental illness through computational analysis of speech. *PLoS One*, *15*(1), e0225695. doi:10.1371/journal.pone.0225695. eCollection 2020.
- Aschbacher, K., O'Donovan, A., Wolkowitz, O. M., Dhabhar, F. S., Su, Y., & Epel, E. (2013). Good stress, bad stress and oxidative stress: insights from anticipatory cortisol reactivity. *Psychoneuroendocrinology*, *38*(9), 1698-1708. doi:10.1016/j.psyneuen.2013.02.004.
- Bhake, R., Kluckner, V., Stassen, H. H., Russell, G. M., Leendertz, J., Stevens, K., Linthorst, A. C. E., & Lightman, S. (2019). Continuous free cortisol profiles - circadian rhythms in healthy men. *Journal of Clinical Endocrinology & Metabolism*, *104*(12), 5935-5947. doi:10.1210/jc.2019-00449.
- Braun, S., Botella, C., Bridler, R., Chmetz, F., Delfino, J. P., Herzig, D., Kluckner, V. J., Mohr, C., Moragrega, I., Schrag, Y., Seifritz, E., Soler, C., & Stassen, H. H. (2014). Affective state and voice: cross-cultural assessment of speaking behaviour and voice sound characteristics. a normative multi-center study of 577+36 healthy subjects. *Psychopathology*, *47*(5), 327-340. doi:10.1159/000363247.
- Braun, S., Annovazzi, C., Botella, C., Bridler, R., Camussi, E., Delfino, J. P., Mohr, C., Moragrega, I., Papagno, C., Pisoni, A., Soler, C., Seifritz, E., & Stassen, H. H. (2016). Assessing chronic stress, coping skills and mood disorders through speech analysis. A self-assessment 'Voice App' for laptops, tablets, and smartphones. *Psychopathology*, *49*(6), 406-419. doi:10.1159/000450959.
- Calear, A.L., & Christensen, H. (2010). Systematic review of school-based prevention and early intervention programs for depression. *Journal of Adolescence*, *33*(3), 429-38. doi:10.1016/j.adolescence.2009.07.004.
- Chopra, K., Katz, J. L., Quilty, L. C., Matthews, S., Ravindran, A., & Levitan, R. D. (2019). Extraversion modulates cortisol responses to acute social stress in chronic major depression. *Psychoneuroendocrinology*, *103*, 316-323. doi:10.1016/j.psyneuen.2019.02.008.
- Cohen, A. S., Renshaw, T. L., Mitchell, K. R., & Kim, Y. (2016). A psychometric investigation of 'macroscopic' speech measures for clinical and psychological science. *Behaviour Research Methods*, *48*(2), 475-486. doi:10.3758/s13428-015-0584-1.
- Corley, L. (2013). Prevalence of mental health issues among college students: how do advisers equip themselves? *The Mentor*, Available from: <http://dus.psu.edu/mentor/2013/08/mental-health-college-students/>
- Cummins, N., Scherer, S., Krajewski, J., Schnieder, S., Epps, J., & Quatieri, T. (2015). A review of depression and suicide risk assessment using speech analysis. *Speech Communication*, *17*, 10-49. doi:10.1016/j.specom.2015.03.004
- Cummins, N., Baird, A., & Schuller, B. W. (2018). Speech analysis for health: Current state-of-the-art and the increasing impact of deep learning. *Methods*, *151*, 41-54. doi:10.1016/j.ymeth.2018.07.007.
- Davidson, M. (2018). The debate regarding maintenance treatment with antipsychotic drugs in schizophrenia. *Dialogues in Clinical Neuroscience*, *20*(3), 2015-2021. doi:10.31887/DCNS.2018.20.3/mdavidson.
- Delfino, J. P., Barragán, E., Botella, C., Braun, S., Camussi, E., Chafrat, V., Mohr, C., Bridler, R., Lott, P., Moragrega, I., Papagno, C., Sanchez, S., Soler, C., Seifritz, E., & Stassen, H. H. (2015). Quantifying Insufficient Coping Behaviour under Chronic Stress. A cross-cultural study of 1,303 students from Italy, Spain, and Argentina. *Psychopathology*, *48*, 230-239. doi:10.1159/000381400.
- Esler, M., Eikelis, N., Schlaich, M., Lambert, G., Alvarenga, M., Dawood, T., Kaye, D., Barton, D., Pier, C., Guo, L., Brenchley, C., Jennings, G., & Lambert, E. (2008). Chronic mental stress is a cause of essential hypertension: presence of biological markers of stress. *Clinical and Experimental Pharmacology and Physiology*, *35*(4), 498-502. doi:10.1111/j.1440-1681.2008.04904.x.
- Faurholt-Jepsen, M., Busk, J., Frost, M., Vinberg, M., Christensen, E. M., Winther, O., Bardram, J. E., & Kessing, L. V. (2016). Voice analysis as an objective state marker in bipolar disorder. *Translational Psychiatry*, *6*, e856. doi:10.1038/tp.2016.123.
- Garrett, K. L., & Healey, E. C. (1987). An acoustic analysis of fluctuations in the voices of normal adult speakers across three times of day. *Journal of Acoustical Society of America*, *82*, 58-62. doi:10.1121/1.395437.
- Grabowski, K., Rynkiewicz, A., Lassalle, A., Baron-Cohen, S., Schuller, B., Cummins, N., Baird, A., Podgórska-Bednarz, J., Pieniążek, A., & Łucka, I. (2019). Emotional expression in psychiatric conditions: New technology for clinicians. *Psychiatry and Clinical Neurosciences*, *73*(2), 0-62. doi:10.1111/pcn.12799.
- Hashim, N. W., Wilkes, M., Salomon, R., Meggs, J., & France, D. J. (2017). Evaluation of voice acoustics as predictors of clinical depression scores. *Journal of Voice*, *31*(2), 256.e1-256.e6. doi:10.1016/j.jvoice.2016.06.006.
- Hollien, H., & Darby, J. K. (1979). Acoustic comparisons of psychotic and non-psychotic voices. In: Hollien H., Hollien P. (Eds), *Current issues in the phonetic sciences* (pp. 609-614). Amsterdam: John Benjamins. doi:10.1075/cilt.9.66hol.
- Hunt, J., & Eisenberg, D. (2010). Mental health problems and help-seeking behaviour among college students. *Journal of Adolescents Health*, *46*(1), 3-10. doi:10.1016/j.jadohealth.2009.08.008.
- Johar, S. (2016). Emotion, affect and personality in speech. The bias of language and paralinguistic. Heidelberg: Springer. doi:10.1007/978-3-319-28047-9\_3.
- Kessler, R. C., Amminger, G. P., Aguilar-Gaxiola, S., Alonso, J., Lee, S., & Ustun, T. B. (2007). Age of onset of mental disorders: a review of recent literature. *Current Opinion in Psychiatry*, *20*, 359-364. doi:10.1097/YCO.0b013e32816ebc8c.
- Kraepelin, E. (1921). Manic depressive insanity and paranoia (trans. By M. Barclay). Edinburgh, UK: Livingstone.
- Lee, D.Y., Kim, E., & Choi, M. H. (2015). Technical and clinical aspects of cortisol as a biochemical marker of chronic stress. *BMB Reports*, *48*(4), 209-216. doi:10.5483/bmbrep.2015.48.4.275.
- Li, Y., Lindsey, B. J., Yin, X., & Chen, W. (2012). A comparison of American and Chinese students' perceived stress, coping styles, and health promotion practices. *Journal of Student Affairs Research and Practice*, *49*(2), 211-227. doi:10.1515/jsarp-2012-6298.
- Lortie, C. L., Thibeault, M., Guitton, M. J., & Tremblay, P. (2015). Effects of age on the amplitude, frequency and perceived quality of voice. *Age (Dordrecht)*, *37*(6), 117. doi:10.1007/s11357-015-9854-1.
- Lott, P. R., Guggenbühl, S., Schneeberger, A., Pulver, A. E., & Stassen, H. H. (2002). Linguistic analysis of the speech output

- of schizophrenic, bipolar, and depressive patients. *Psychopathology*, 35(4), 220-227. doi:10.1159/000063831.
- Lötscher, K., Anghelescu, I.G., Braun, S., Bridler, R., & Stassen, H. H. (2010). Polypharmacy in psychiatry: clinical practice versus empirical evidence. *European Neuropsychopharmacology*, 20 (Suppl. 3), 378-379.
- McGinnis, E. W., Anderau, S. P., Hruschak, J., Gurchiek, R. D., Lopez-Duran, N. L., Fitzgerald, K., Rosenblum, K. L., Muzik, M., & McGinnis, R. S. (2019). Giving voice to vulnerable children: machine learning analysis of speech detects anxiety and depression in early childhood. *IEEE Journal of Biomed Health Information*, 23(6), 2294-2301. doi:10.1109/JBHI.2019.2913590.
- Mental Health Foundation. (2016). *Fundamental facts about mental health*. London: Mental Health Foundation.
- Mohr, C., Braun, S., Bridler, R., Chmetz, F., Delfino, J. P., Kluckner, V. J., Lott, P., Schrag, Y., Seifritz, E., & Stassen, H. H. (2014). Insufficient coping behaviour under chronic stress and vulnerability to psychiatric disorders. *Psychopathology*, 47, 235-243. doi:10.1159/000356398.
- Oswald, L. M., Zandi, P., Nestadt, G., Potash, J. B., Kalaydjian, A. E., & Wand, G. S. (2006). Relationship between cortisol response to stress and personality. *Neuropsychopharmacology*, 31, 1583-1591. doi:10.1038/sj.npp.1301012.
- Pigott, H. E., Leventhal, A. M., Alter, G. S., & Boren, J. J. (2010). Efficacy and effectiveness of antidepressants: current status of research. *Psychotherapy and Psychosomatics*, 79(5), 267-279. doi:10.1159/000318293.
- Püschel, J., Stassen, H. H., Bomben, G., Scharfetter, C., & Hell, D. (1998). Speaking behaviour and voice sound characteristics in acute schizophrenia. *Journal of Psychiatric Research*, 32, 89-97. doi:10.1016/S0022-3956(98)00046-6.
- Safer, D. J. (2019). Overprescribed medication medications for US adults: four major examples. *Journal of Clinical Medicine Research*, 11(9), 617-622. doi:10.14740/jocmr3906.
- Schuch, F. B., Vancampfort, D., Firth, J., Rosenbaum, S., Ward, P. B., Silva, E. S., Hallgren, M., Ponce De Leon, A., Dunn, A. L., Deslandes, A. C., Fleck, M. P., Carvalho, A. F., & Stubbs, B. (2018). Physical activity and incident depression: a meta-analysis of prospective cohort studies. *American Journal of Psychiatry*, 175(7), 631-648. doi:10.1176/appi.ajp.2018.17111194.
- Sila-Nowicka, K., & Thakuriah, P. (2019). Multi-sensor movement analysis for transport safety and health applications. *PLoS One*, 14(1), e0210090. doi:10.1371/journal.pone.0210090.
- Slavich, G. M., Taylor, S., & Picard, R. W. (2019). Stress measurement using speech: Recent advancements, validation issues, and ethical and privacy considerations. *Stress*, 22(4), 408-413. doi:10.1080/10253890.2019.1584180.
- Stallman, H. M., & Shochet, I. M. (2009). Prevalence of mental health problems in Australian university health services. *Australian Psychologist*, 44(2), 122-127. doi:10.1080/00050060902733727
- Stallman, H. M. (2010). Psychological distress in university students: A comparison with general population data. *Australian Psychologist*, 45(4), 249-257. doi:10.1080/00050067.2010.482109.
- Stassen, H. H., Kuny, S., & Hell, D. (1998). The speech analysis approach to determining onset of improvement under antidepressants. *European Neuropsychopharmacology*, 8(4), 303-310. doi:10.1016/s0924-977x(97)00090-4.
- Stassen, H. H., Angst, J., Hell, D., Scharfetter, C., & Szegedi, A. (2007). Is there a common resilience mechanism underlying antidepressant drug response? Evidence from 2848 patients. *Journal of Clinical Psychiatry*, 68(8), 1195-1205. doi:10.4088/jcp.v68n0805.
- Stassen, H. H., Anghelescu, I.G., Angst, J., Böker, H., Lötscher, K., Rujescu, D., Szegedi, A., & Scharfetter, C. (2011). Predicting response to psychopharmacological treatment. Survey of recent results. *Pharmacopsychiatry*, 44, 263-272. doi:10.1055/s-0031-1286290.
- Stassen, H. H., Bachmann, S., Bridler, R., Cattapan, K., Herzig, D., Schneeberger, A., & Seifritz, E. (2021). Inflammatory processes linked to major depression and schizophrenic disorders and the effects of polypharmacy in psychiatry: Evidence from a longitudinal study of 279 patients under therapy. *European Archives of Psychiatry and Clinical Neurosciences*, 271, 507-520. doi:10.1007/s00406-020-01169-0.
- Staufenbiel, S. M., Penninx, B. W., Spijker, A. T., Elzinga, B. M., & van Rossum, E. F. (2013). Hair cortisol, stress exposure, and mental health in humans: a systematic review. *Psychoneuroendocrinology*, 38, 1220-1235. doi:10.1016/j.psyneuen.2012.11.015.
- Taguchi, T., Tachikawa, H., Nemoto, K., Suzuki, M., Nagano, T., Tachibana, R., Nishimura, M., & Arai, T. (2018). Major depressive disorder discrimination using vocal acoustic features. *Journal of Affective Disorders*, 225, 214-220. doi:10.1016/j.jad.2017.08.038.
- Titze, I. R. (1994). *Principles of voice production*. Englewood Cliffs, NJ: Prentice Hall.
- Wang, J., Zhang, L., Liu, T., Pan, W., Hu, B., & Zhu, T. (2019). Acoustic differences between healthy and depressed people: a cross-situation study. *BMC Psychiatry*, 19(1), 300. doi:10.1186/s12888-019-2300-7.
- Zhang, M., Bridler, R., Mohr, C., Moragrega, I., Sun, N., Xu, Z., Yang, Z., Possenti, M., & Stassen, H. H. (2019). Early detection of the risk of developing psychiatric disorders: a study of 461 Chinese University Students under chronic stress. *Psychopathology*, 52(6), 367-377. doi:10.1159/000505787.