



WHITEPAPER

Voice: An Indicator of Stress

Health & Wellness Case Study

VOICE: AN INDICATOR OF STRESS

Abstract / A group of 340 individuals came together to test whether voice could be a strong indicator of stress compared to the current gold standard - self reporting questionnaires. The results of this exercise established a foundation for further groundbreaking work in mental health.

Introduction / Are you stressed? Self-reporting one's mental health has its advantages and disadvantages. In many cases it is the only available indicator of how an individual is feeling at the moment. What if there was a non-invasive, objective way to gauge a person's level of stress or depression? As humans we use voice to signal emotion. We can tell, just from voice, that an individual is sad, mad, or scared. Just as we do, a computer using AI (Artificial Intelligence) and machine learning can determine if a person is anxious.

How does it work? / Canary Speech has an internationally patented approach to using voice to identify disease. Our scientists have discovered 2,548 biomarkers such as, tonal quality and length of pauses between words that are used to create mathematical equations. These models can then be used to evaluate disease from conversational speech in a matter of seconds.

Exercise / Of the 340 users who participated, 297 completed all self assessments and all audio responses giving an 87 percent participation compliance rate. The exercise began with five gold standard assessments - STAI (Anxiety), PANAS (Mood), PH-9 (Depression), PSQI (Sleep Quality) GAD7 (Anxiety).



The participants then moved to the audio collection portion of the exercise where they read ten sentences. They then proceeded to read a short paragraph. The voice assessment ended with the audio collection of free speech with a prompt given, “How are you feeling?” “What made you feel that way?” All participants used their own smart phone to complete the exercise.

The Results / In general, the **score distributions matched the distribution of the general population**, with the slight exception of a spike in max value for GAD7 and PHQ9, possibly caused by inaccurate self assessments.

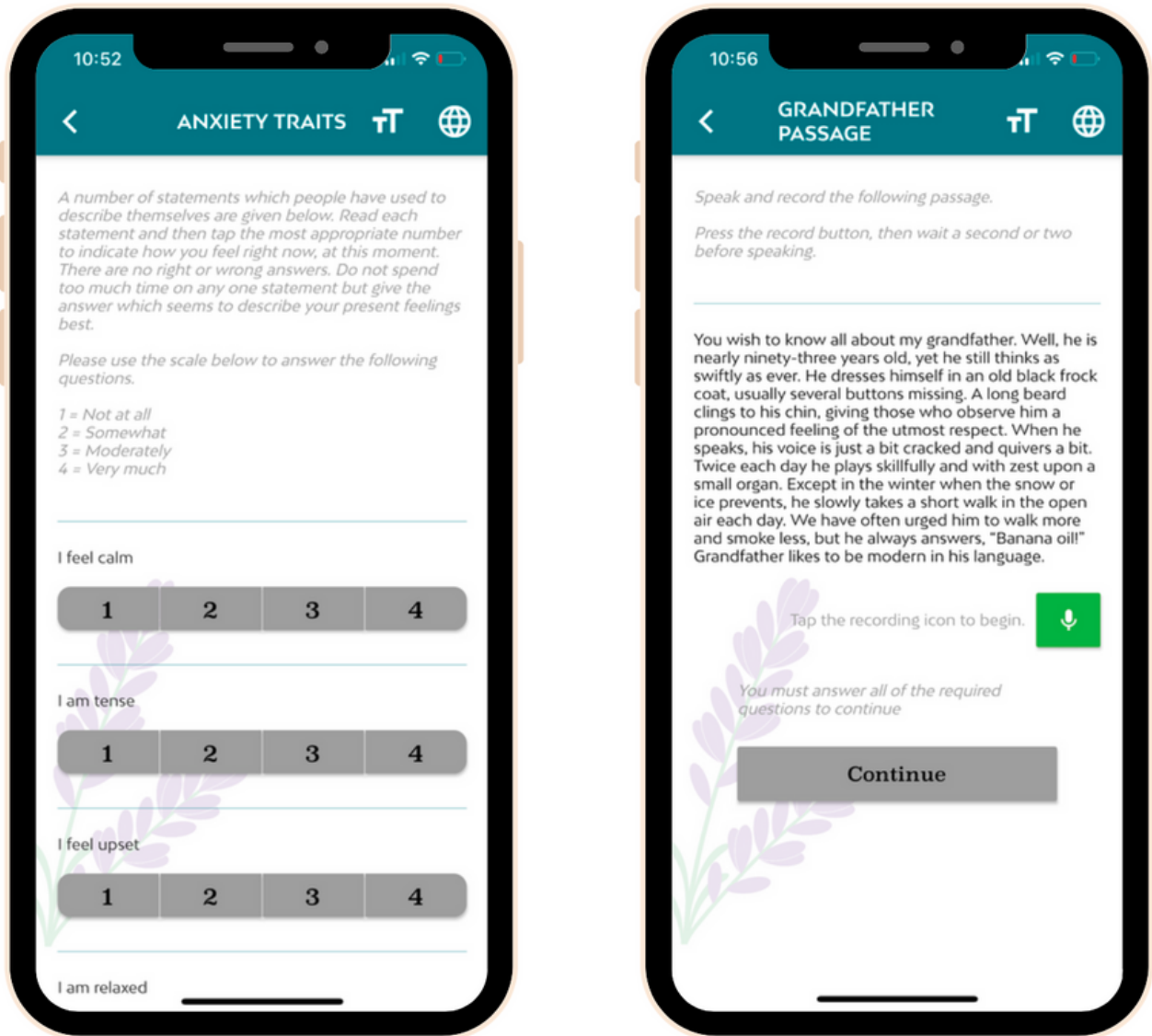
Around **80 percent of users** successfully responded for a full minute with the remaining finishing within a shorter time.

The type of voice features fell in two categories: acoustic and linguistic. Acoustic features capture signal-level modulations due to the speaker's status, while linguistic features capture language-level patterns which may be influenced by the condition. Acoustic features are calculated on a per-frame basis. Frames are defined as 25 ms sliding windows that are created every 10 ms. This resulted in **33,384 features being extracted and correlated**. The performance was measured in terms of Pearson’s correlation coefficients (CC) between VQ scores and predicted scores.

In conclusion, the voice sample correlated independently alongside the self-reporting assessments, giving a clear indication of the case for voice as an objective and non-invasive mental health evaluation tool.

Furthermore, the vocal assessments (orange) correlated with the national average scores of the gold standard questionnaires (blue), such as the STAI and GAD7, giving a clear indicator that stress and mood can successfully be measured by vocal models using AI and machine learning.

CANARY SPEECH



	STAI	GAD7	PSQI	PANAS	PHQ9
Lang + Acoustic	0.97	0.92	0.98	0.88	0.92

