# Chatbot-based assessment of employees' mental health: design process and pilot implementation

Article (Accepted Version)

Original Paper

# A Chatbot-based Assessment of Employees' Mental Health: Design Process and Pilot Implementation

Ines Hungerbuehler[1], Kate Daley[1], Kate Cavanagh[2], Heloisa Garcia Claro[3,4], Michael Kapps[1]

[1]Vitalk, TNH Health, São Paulo, Brazil

[2] School of Psychology, University of Sussex, Brighton, UK

[3]Department of Preventive Medicine, University of São Paulo Medical School, São Paulo, Brazil

[4]School of Nursing, University of Campinas, Campinas, São Paulo, Brazil

## Abstract

*Background:* Stress, burnout and mental health problems, such as depression and anxiety are common and can significantly impact on workplaces through absenteeism and reduced productivity. To address this issue, organizations must first understand the extent of the difficulties by mapping the mental health of their workforce. Online surveys are a cost-effective and scalable way to do this but typically have low response rates, in part due to a lack of interactivity. Chatbots offer one potential solution, enhancing engagement through simulated natural human conversation and use of interactive features. *Objective:* To explore if a text-based chatbot is a feasible way to engage and motivate employees to complete a workplace mental health assessment. The paper describes the design process and results of this pilot implementation. *Methods:* A fully automated chatbot ('Viki') was developed to evaluate employee risks of suffering from depression, anxiety, stress, insomnia, burnout and work-related stress. Viki uses a conversation style and gamification features to enhance engagement. A cross-sectional analysis has been conducted in order to gain first insights of a pilot implementation within a small to medium-sized enterprise (N=120). *Results*: The response rate was 64%. In total, 98 employees started the assessment, and 77 (79%) completed it. The majority of people scored in the mild range for anxiety (20/40; 50%) and depression (16/28; 57%), moderate range for stress (10/22, 46%) and subthreshold level for insomnia (14/20, 70%) as defined by their questionnaire scores. *Conclusions*: A chatbot-based workplace mental health assessment seems to be a highly engaging and effective way to collect anonymized mental health data among employees with response rates comparable to face-to-face interviews.

*Keywords:* Chatbot; conversational agent; online; digital health; mobile phone; mental health; workplace; work stress; survey; response rate

## Introduction

We spend, on average, one-third of our life at work. Thus, the workplace is one of the key environments that can affect our quality of life, emotional and physical wellbeing. Generally, work is considered good for our mental health with involuntary joblessness being a well-recognized risk factor for mental health problems, including depression [1]. However, stressful work conditions can also contribute to the development of mental health problems [2]. Causes of common mental health problems such as depression and anxiety are complex and may include traumatic life experiences or inherited traits, but can also be a reaction to work-related stress. The Dunedin Study found high-demand jobs were associated with onset of depression and anxiety in people with no prior history of diagnosis or treatment for either disorder [3]. Generally, work-related stress is considered a consequence of the organization and management, the skills and competencies of employees, and the support they receive [2].

Besides serious consequences for the individuals' mental health and the direct medical costs, stressful working conditions have indirect costs though reduced productivity due to absenteeism and presenteeism. A recent WHO-led study estimates that depression and anxiety disorders cost the global economy US$ 1 trillion each year in lost productivity [4]. The implementation of cost effective and feasible interventions could therefore have a significant impact on the individual, the organization and the economy [5].

In order to minimize the impact of workplace risk factors, adequate policies and intervention programs should be implemented. To determine the appropriate interventions, current mental health of employees and possible sources of work-related stress must first be assessed. There are specific measures that are commonly used to evaluate different mental health conditions and work-related risk factors such as the Patient Health Questionnaire [6], Depression, Anxiety and Stress Scale [7], and Job Satisfaction Survey [8]. Traditionally these are self-administered in person or over the telephone with a health professional who facilitates completion, offering encouragement or clarification as needed. Both approaches yield similar results [9]. It is also possible to complete the scales online, using a webpage or smartphone app to display questions and collect user responses. This method has been found to be feasible in both clinical [10] and workplace settings [11] albeit with relatively low response rates (34%) [11]. Symptom scores reported via smartphone app strongly correlate with those reported through traditional paper and pen methods, although symptoms reported via smartphone were on average 3 points higher [12].

Engagement varies widely among digital health programs and smartphone apps [13]. According to a recent review, low engagement can occur when apps are not designed with the user in mind or do not solve the problem the user cares most about [14]. A lack of interactive or engaging features can also increase the risk of survey fatigue where users get tired and do not finish the survey [14]. Guided self-help interventions are found to have greater adherence than non-guided interventions [15]. This suggests that human support could improve engagement, however this limits scalability. Chatbot-driven conversational surveys could offer an alternative by automating this encouragement and interaction. Chabots have been found to have significantly greater engagement and higher quality responses than typical online surveys [16,17] and the relative anonymity offered by such an approach could be of additional value to employees.

Chatbots, or conversational computer programs, simulate human conversation. Features include word-classification processes, natural language processing and artificial intelligence in addition to simple keywords scans and databases linking common phrases and predefined responses which help the chatbot tailor answers to a specific user input. Most chatbots are accessed via websites or mobile applications, or can be integrated into virtual assistants as a conversational component of a system, which can also control external devices or manage basic tasks such as emails or to-do lists for example. Chatbots tend to be represented by an animated character, in some cases an embodied ‘human’ conversational agent who uses and responds to verbal and nonverbal communication such as hand gestures or body posture.

Many interactions between organizations and customers are already bot-driven, enabling companies to respond to more people at a faster and cheaper rate than if they use human customer service representatives. Besides being a cost-effective and feasible way to communicate, chatbots are capable of generating a believable and dynamic dialogue. This has the potential to enhance engagement rates, using the chatbot to successfully guide and motivate users. Compared to typical online surveys, this may result in a higher level of engagement [16], and greater symptom disclosure [17]. Nothing of this nature currently exists for workplace mental health in Brazil as far as the authors are aware.

This paper describes the chatbot design process and results of a pilot implementation in a workplace setting. The paper seeks to explore if a text-based chatbot is a feasible way to engage and motivate employees to complete a workplace mental health assessment, and so provide important insights for the employee and organization.

# Methods

## Design

The present article describes the results and insights of a pilot implementation of a chatbot-based mental health assessment in a real-world workplace setting, based on a cross-sectional analysis.

## Sample

The sample of the present analysis are employees of an industrial plant in São Paulo, Brazil, with a total of 120 employees, who participated in the assessment between October and November 2019. Those employees work on the recycling plant (n=52), in reverse logistics operations (n=40) or in the office in administration, information technology or human resources roles (n=28).

## Approval

Consent was given by participants as part of the onboarding process. They agreed for their anonymized data to be used for research purposes and aggregated data to be shared with the organization. Data was stored securely and password protected.

The company board of directors approved the use of the routinely collected data for analysis and publication after having been briefed about the assessments and privacy requirements.

Research based on aggregate user data with no possibility of individual identification does not require approval of the Research Ethics Committee (CEP/CONEP) in Brazil.

## Chatbot development and testing

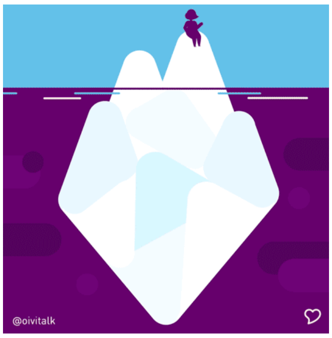
The chatbot was designed to assess employees' mental health in a cost-effective and engaging way. This requires a specific developing platform, clinically validated content and an adequate visual presentation, but also a clear purpose and a well defined personality. A key component for reaching this goal is the user experience (UX), used to help connect the chabot with users and build a shared experience. The first step was to analyze the needs, characteristics, and behaviours of the target user group. Based on those insights, the chatbot avatar Viki (Figure 1), her language and conversation style were created. The UX team conducted focus groups and telephone interviews with potential users to verify if the chatbot fits with their expectations and needs. They were asked, for example, if the objective was clear, if the length of the check-up was accurate, whether they had any difficulties in completing the check-up, how they rated their experience of communicating with the chatbot and how they would describe Viki. The data obtained from these interviews were used to refine the chatbot.

Figure 1: The chatbot avatar Viki

The chatbot is built on ruby and javascript and was created by the team at TNH Health. The check-up assessment is rule-based, with next steps determined by user responses. As part of the design process, decision trees were defined and all possible user journeys mapped and analyzed. Decision trees help the chatbot provide the right responses and information based on the users inputs, that is to say, to customize the conversation. If the user, for example, responds that he/she wants to know more about depression, the chatbot delivers further information about the topic. If the user prefers to continue without knowing more about it, the chatbot takes him/her to the next topic. Building a decision tree creation tool with all of the necessary settings for interactions and data organization allowed the chatbot to be updated in an agile way without needing a new system release, something which was crucial to the development.

Engagement was known to be challenging, and so gamification features were added to address this issue. Gamification is defined as “the use of game design elements in nongame contexts” [18] and features include levels, challenges, points, progress, feedback, story and reward [19]. The most important features utilized here were story and feedback. For the story, Viki guides users through the assessment process, presented as an expedition around an iceberg (Figure 2). The iceberg represents issues relating to mental health and is divided into seven different sections: stress, anxiety, depression, burnout and work-related stress. Insomnia is additionally presented if the user scores positive to the question “I had trouble falling or staying asleep or sleeping too much in the depression block.

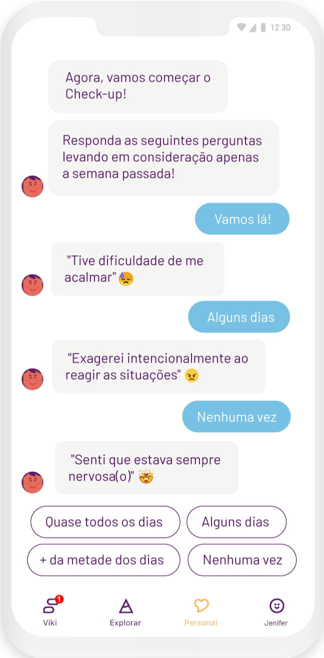
Figure 2: The iceberg story




During the conversations, more and more of the iceberg, and thus the key topics become visible to the user. Each section consists of an introduction to the topic and a standardized questionnaire. The questionnaires are delivered in conversation format, with Viki asking a question and the user selecting their answers from predefined responses, using the standard options for each questionnaire (Figure 3). Only once the user selects their answer is the next question presented. This is on a rolling screen with approximately three responses visible at one time. The user can go back and change their answer if they want, but there was no functionality to skip a question. During the check-up, besides the questions of the different scales, Viki also offers messages of encouragement. Those messages are designed to keep the participant motivated and engaged, the same way as a real human interviewer would do. An example of such a message is “We are almost done, you are doing very well!”. All communication is text-based, written in Brazilian Portuguese.

The whole assessment takes approximately 15 minutes. Responses are captured automatically and stored within a secure database. The user issue of non-response was discussed in the design process and the team agreed that a limit of 24 hours should be set. The user can pause at any time during the assessment, but if they do not complete within 24 hours the system will reset and data is overwritten. This ensures that all questionnaire data has been collected within a specific time period.

Figure 3: Screenshot example of check-up




Immediately after completing the assessment, participants receive personalized feedback and recommendations. Participants are reminded that the results do not offer a diagnosis, but may indicate the presence of an emotional problem. In the case of a serious risk being identified (very high risk for anxiety or depression, and/or suicide risk detected during the assessment, or via key words typed by users) the conversation follows a safety protocol which includes referral to the care network or, if necessary, to emergency services.

Aggregated and anonymized data are presented in a dashboard and organization report to the company's management. This allows oversight of the assessment process in real-time (response rate, distributions of mental health outcome categories of already assessed employees, etc.) and offers valuable insights to the organization, identifying issues and recommending actions, for example, further campaign actions in case of a general low engagement of the employees or planning of target-oriented interventions for specific departments or positions based on the mental health outcome distributions.
Participants are informed within the terms of use, which they are asked to agree at the beginning of the check-up, about the anonymized sharing of their data with the company. In order to preserve anonymity, no name or other identification date are included on the dashboard or report and departments or sections of less than eight persons are pooled into bigger groups.

The usability and technical functionality of the chatbot and dashboards were tested prior to launch with users (with internal staff and organization employees). The only issue raised during this testing was how the chatbot would react to an unexpected free text entered by the employee during the assessment. Here the settings and natural language understanding (NLU) were adjusted to enable smooth running of the chatbot whilst adhering to risk protocols.

## Mental health outcomes

The check-up contains the following questionnaires to cover topics of anxiety, depression, stress, burnout, and work-related stress. Insomnia was added for users who did not answer the question about difficulties with sleeping with "Not at all" in the previously applied Patient Health Questionnaire (PHQ-9). All questionnaires have been translated and validated in Brazilian Portuguese, with good psychometric properties [20-25]. No changes were made to question order or wording.

The General Anxiety Disorder (GAD-7) scale is a 7-item self report scale used to assess anxiety symptoms over the past two weeks *(e.g. how often have you been bothered by feeling afraid something awful might happen)*. Responses range from 0 (not at all) to 3 (nearly everyday). Total scores are divided into four categories: none (0-4), mild (5-9), moderate (10-14) and severe (15+) symptoms [26].

The Patient Health Questionnaire-9 (PHQ-9) is a 9 item self report scale that evaluates symptoms of depression over the past two weeks *(e.g. how often have you been bothered by feeling down, depressed, or hopeless)*. Responses range from 0 (not at all) to 3 (nearly every day). Total scores are divided into five categories: none (0-4), mild (5-9), moderate (10-14), moderately severe (15-19) and severe (20+) symptoms [6].

The Depression, Anxiety & Stress Scale (DASS-21) is a 21 item self report questionnaire consisting of three scales to measure depression, anxiety and stress [27]. The stress subscale consists of 7 items and the user is asked how much each statement applied to them in the past week *(e.g. I found it difficult to relax)*. Responses range from 0 (did not apply) to 3 (applied very much or most of the time). Total scores are doubled and divided into five categories: normal (0-14), mild (15-18), moderate (19-25), severe (26-33), extremely severe (34+) levels of stress [7].

The Insomnia Severity Index (ISI) is a 7-item self-report questionnaire assessing the nature, severity, and impact of insomnia. Responses range from 0 (no problem) to 4 (very severe problem). Total scores are categorized as: absence (0–7), sub-threshold (8–14), moderate (15–21), and severe insomnia (22–28) [28].

The Oldenburg Burnout Inventory (OLBI) is a 16 item self report scale with two dimensions, exhaustion and disengagement from work. *(e.g. there are days when I feel tired before I arrive at work.)*. Answers range from from 1 (strongly agree) to 4 (strongly disagree) [29]. For the analysis of the present results, the following categories were used: very low (0-15), low (16-30), high (31-45) and very high risk of burnout (46+).

The Job Stress Scale (JSS) questionnaire is based on a two-dimensional theoretical model created by Robert Karasek that relates two aspects at work, demands and control, to the risk of illness. A third dimension, support, was later added by Theorell as a protective factor to guard against work-related stress. A 15 item short form of the questionnaire was used to assess work-related risk and protective factors *(e.g. does your work demand too much effort? Do you have a choice in deciding how you do your work?).* Scores range from 1 (almost never / totally agree) to 4 (frequently / totally disagree) [8]. Categories used were: normal (0-15), slightly increased (16-30), increased (31-45) and extremely increased (46+).

### Recruitment and implementation

As the first stage in implementation the mental health assessment was presented to the managers of each department of the company. After that, an internal multi-channel information campaign was used to present the assessment to employees and educate them about the importance of mental health. Viki was introduced as part of this campaign, using a variety of online and offline channels (email, intranet, banners and leaflets). This psychoeducative action aimed to reduce stigma and motivate employees to consider looking at their own mental health, as well as to generate trust in the product.

All employees were then emailed a weblink with an invitation to complete the check-up to obtain feedback on their scores. No financial incentive was offered and employees were informed the check-up was anonymous and voluntary. They were told it would take up to 15 minutes, and no stipulations were made regarding whether to complete this at work or home. It could be completed via cell phone or computer.

When employees click the link, they first pass through an authentication process where they register and enter their company code (which helped to identify the participants of this specific company survey within the whole user population) and agree to the terms and conditions. Viki then begins the check-up.

### Statistical Analysis

To describe sample characteristics, baseline symptoms and completion rates, descriptive statistics were used. Only complete data sets were included in this analysis, and duplicates were not possible due to the registration process used. The Checklist for Reporting Results of Internet E-surveys was used for data reporting [30].

## Results

### Sample and engagement

Of the 120 employees approached to take part, 77 (79%) completed it. This response rate enabled data to be obtained for 64% of the organization's workforce. The data collection was conducted within four weeks. More than half of those responding did so within the first week of the campaign. Of the 98 employees who started the assessment, 79% completed it (completion rate).

Of the 77 employees completing the check-up, the majority were male (45; 58%), white (40; 52%), and aged between 25 and 44 years (55; 71%), reflecting the demographics of the organization's workforce. The majority of respondents were in occupations of ‘plant and machine operators and assemblers’ (n=21, 27%), or ‘technicians and associate professionals’ (n=19, 25%).

## Implementation

During implementation we noticed differences in speed of roll out, and considered if the motivation and conviction at a management level could be a factor in engagement. As a preventative measure to maximise engagement we contacted department managers to emphasise the benefits of the check up, and to troubleshoot any potential issues. No issues were identified but some managers reported they would ensure they cascade this information to their teams. The informal feedback received from the lead of the project was that the campaign had been well received and they felt the chatbot added value. This evidence is of course more anecdotal as it did not form part of the data analysis.

## Mental health risks

Overall, the sample scored in the low ranges on the majority of the questionnaires. Scores were higher in the area of work-related stress, with equal control and demand reported alongside lower support. On the PHQ-9 44% (34/77) of users reported difficulties with sleep and so completed the ISI for insomnia.

Table 1: Means and standard deviations for each questionnaire measure

| | n | Mean (SD) | Category |
|---|---|---|---|
| Anxiety (GAD-7) | 77 | 6.21 (4.56) | Low |
| Depression (PHQ-9) | 77 | 4.40 (5.21) | None |
| Stress (DASS-21) | 77 | 11.09 (7.13) | Normal |
| Insomnia (ISI)* | 34 | 9.26 (5.66) | Sub threshold |
| Burnout (OBLI) | 77 | 27.68 (8.38) | Low |
| Occupational stress (JSS)<br>- Control<br>- Demand<br>- Support | 77 | 32.38 (3.55)<br>12.32 (1.99)<br>12.19 (1.72)<br>7.86 (2.40) | Increased |
| *completed only if sleep was identified as an issue | | | |

In terms of those with symptoms, the majority of people scored in the mild range for anxiety (20/40; 50%), the mild range for depression (16/28; 57%), subthreshold level for insomnia (14/20, 70%), and moderate level for stress (10/22, 46%) as defined by the questionnaire scores. For burnout, most scored in the low risk category (50/74, 68%) and increased risk category for job-related stress (53/77, 69%).

## Discussion

The chatbot-based assessment was successfully implemented in the workplace, suggesting a chatbot could be a feasible way to engage employees in completing a workplace mental health assessment. There was a 79% (77/98) completion rate, obtaining questionnaire responses from 64% (77/120) of the workforce. This compares favourably to face-to-face data collection methods, for example the São Paulo Megacity study which reported a response rate of 81% [31] and epidemiologic studies which reported a response rate of 70% using the same method of data collection [32]. The completion rate in this analysis is 25% higher than that of results found in a telecommunication company using a web-based screening for depression [11] and also clearly higher than that of other online surveys where rates less than 10% are not uncommon [33]. A previous study reported a similar completion rate (78%) using a smartphone app to administer up to three survey sessions per day in a sample of psychiatric outpatients. However this was with a smaller sample and participants were financially compensated which may influence motivation [12]. It would be interesting to explore the impact of the different components of implementation to ascertain the factors that are most integral for success whether this be the chatbot or the onboarding process. An online tool and distribution method could be particularly pertinent in an era of remote working.

As shown in the results, the majority of users obtained low scores on all of the questionnaires. However, many users did report symptoms at a moderate and severe level. The proportion scoring at this level for depression (20/77, 26%) and anxiety (12/77, 16%) are higher than estimated prevalence rates of anxiety disorders (9.3%) and depressive disorders (5.8%) found in the general population in Brazil [34], although rates do vary considerably depending on method of data collection and measure used [35]. Rates are normally higher in women than in men [35] which is interesting in this sample given it was predominately male. The risk of burnout is similar to levels found in other professional groups such as health workers [36] and teachers [37]. High demand was often reported in combination with a low degree of perceived control, both risk factors for burnout. This combination is considered the most critical in terms of a negative impact on individual mental health. Presence of anxiety, depression and insomnia could also impact on organizational productivity through absenteeism and presenteeism [2,11] which requires further exploration. The fact that some people reported symptoms shows the importance of addressing mental health issues within the workplace. It would be interesting to know whether these people have ever sought professional help or support for these difficulties. Insights gained from the assessment could be used to identify individual and organizational level strategies that could be implemented to improve mental health and potentially, productivity within the workplace.

## Limitations

The questionnaire measures have not yet been validated for use in a chatbot format with gamification. As the pilot has shown implementation to be feasible, the next step is to complete a validation study to assess the effect of using this method on the psychometrics properties of the questionnaires. Establishing validity is required before conclusions can really be drawn regarding the mental health of the workplace.

Whilst the chatbot was successfully able to obtain data for 64% of the workforce, there is no data on those who did not participate or those who dropped out during the assessment. It would be interesting to compare socio-demographics and baseline symptoms between groups. Previous research indicates for example that being male, with lower educational level and comorbidity of depression with anxiety can increase risk of dropout and non-engagement [38]. Insights could be used to adapt the chatbot or onboarding process to make it more appealing to the less-engaged group. Taking age and gender into account has been found to enhance use of e-mental health programs [39].

As this was a cross-sectional data analysis, we do not know if the high response rate will be sustained over time. There is the possibility that the response rate was inflated due to the novelty of the approach, something that has been suggested of mHealth interventions [40]. Longitudinal research could explore this issue and could assist in the development of UX with chatbots over time. Additionally, as the system used for output generation (responses to the users input) is fixed (based on predefined decision trees), the conversation opportunities here are limited. A larger amount of conversational data gained over time will be necessary to train the chatbot and make it ~~more~~ intelligent in a more autonomous way. This would also allow free text interpretation, for example.

Without controls it is unclear if results are due to the specific sample, their workplace or the method of data collection which may facilitate higher levels of disclosure [11]. After validating the questionnaires for use in this format, it would be important to repeat this exercise with different workplaces to ascertain if results are generalizable to other populations, particularly considering issues of gender, age, work type and level of education. It would be important to replicate the implementation in a workplace with more gender balance, and larger sample size. A usability questionnaire would further strengthen the validity of the results.

Further research is planned which will include comparison with traditional paper and pen methods, and web-based forms. These studies will also include repeat measurement following the implementation of remedial measures to ascertain their impact. We would also like to explore, using implementation science, the factors and processes involved in successful implementation of the chatbot within the workplace since there are many variables involved.

## Conclusions

The creation of healthy workplaces and adequate mental health policies need to be based on a comprehensive needs assessment. Face to face assessments are not anonymous or scalable, and online surveys are often limited by low response rates. A chatbot offers a fully automated digital solution, incorporating gamification features to engage and motivate employees to complete a workplace mental health assessment. The Vitalk chatbot was found to have response rates comparable to face-to-face interviews, suggesting this could be a feasible way to collect this data. In order to further verify this new solution, a validation study comparing it with other formats, such as face-to-face interviews or online surveys, as well as including a feasibility and satisfaction analysis, would be the logical next step.

## Acknowledgements

Thank you to Raphael Mota and Bruno de Castro Paul Schultze for assisting with the additional data extraction and analysis, to Aimee da Silva Ferreira and Carolina Nakatsu for the graphic design.

IH made substantial contributions to the conception and design of the work and completed analysis or results. All authors contributed to the writing, review and approval of this manuscript for publication.

## Conflicts of Interest

IH, KD and MK work for TNH Health, the company who developed the chatbot and provided funding for this publication. KC and HC declare to have no conflicts of interest.

## Abbreviations

| | |
|---|---|
| DASS-21 | Depression, Anxiety and Stress Scale-21 |
| GAD-7 | Generalized Anxiety Disorder-7 |
| HPQ | World Health Organization Health and Work Performance Questionnaire |
| ISI | Insomnia Severity Index |
| JSS | Job Stress Scale |
| MBI | Maslach Burnout Inventory |
| OLBI | Oldenburg Burnout Inventory |
| PHQ -9 | Patient Health Questionnaire-9 |
| SME | Small and Medium Enterprise |
| UX | User Experience |
| WHO | World Health Organization |

## References

1. Jefferis BJ, Nazareth I, Marston L, Moreno-Kustner B, Bellón JA, Svab I, Rotar D, Geerlings MI, Xavier M, Goncalves-Pereira M, Vicente B, Saldivia S, Aluoja A, Kalda R, King M. Associations between unemployment and major depressive disorder: Evidence from an international, prospective study (the predict cohort), Social Science & Medicine; 2011, 73(11): 1627-1634. https://doi.org/10.1016/j.socscimed.2011.09.029. PMID: 22019370
2. World Health Organisation. Mental Health in the Workplace. Information Sheet. [Last accessed on 11 February 2020]. Available from: http://www.who.int/mental_health/in_the_workplace/en/
3. Melchior M, Caspi A, Milne BJ, Danese A, Poulton R, Moffitt TE. Work stress precipitates depression and anxiety in young, working women and men. Psychol Med; 2007, 37(8):1119-1129. https://doi.org/10.1017/S0033291707000414. PMID: 17407618
4. Chisholm D, Sweeny K, Sheehan P, Rasmussen B, Smit F, Cuijpers P, Saxena S. Scaling-up treatment of depression and anxiety: a global return on investment analysis. www.thelancet.com/psychiatry Published online April 12, 2016. 10.1016/S2215-0366(16)30024-4 PMID: 27083119 DOI:
5. Bertram M, Banatvala N, Kulikov A, Belausteguigoitia I, Sandoval R, Hennis A. Using economic evidence to support policy decisions to fund interventions for non-communicable diseases. BMJ; 2019; 365:l1648. https://doi.org/10.1136/bmj.l1648. PMID: 31110025.
6. Kroenke K, Spitzer RL.The PHQ-9: A New Depression Diagnostic and Severity Measure. Psychiatric Annals; 2002, 32(9):1-7. https://doi.org/10.3928/0048-5713-20020901-06
7. Antony MM, Bieling PJ, Cox BJ, Enns MW, Swinson RP. Psychometric properties of the 42-item and 21-item versions of the depression anxiety stress scales in clinical groups and a community sample. Psychological Assessment; 1998, 10:176-181. https://doi.org/10.1037/1040-3590.10.2.176
8. Theorell T. The demand-control-support model for studying health in relation to the work environment: an interactive model. In: Orth-Gómer K, Schneiderman N, editors. Behavioral medicine approaches to cardiovascular disease. Mahwah, NJ: Lawrence Erlbaum Associates; 1996, 69-8.
9. Pinto‐Meza A, Serrano‐Blanco A, Peñarrubia MT, Blanco E, Haro JM. Assessing Depression in Primary Care with the PHQ‐9: Can It Be Carried Out over the Telephone? Journal of General Internal Medicine; 2005, 20(8):738-742. https://doi.org/10.1111/j.1525-1497.2005.0144.x. PMID: 16050884.
10. Fann JR, Berry DL, Wolpin S, Austin‐Seymour M, Bush N, Halpenny B, Lober WB, McCorkle R. Depression screening using the Patient Health Questionnaire‐9 administered on a touch screen computer. Psycho-oncology; 2009, 18(1):14-22. https://doi.org/10.1002/pon.1368. PMID: 18457335
11. Harvey SB, Glozier N, Henderson M, Allaway S, Litchfield P, Holland-Elliott K, Hotopf M. Depression and work performance: an ecological study using web-based screening, Occupational Medicine; 2011, 61(3): 209-211. https://doi.org/10.1093/occmed/kqr020. PMID: 21525074.

12. Torous J, Staples P, Shanahan M, Lin C, Peck P, Keshavan M, Onnela JP. Utilizing a Personal Smartphone Custom App to Assess the Patient Health Questionnaire-9 (PHQ-9) Depressive Symptoms in Patients With Major Depressive Disorder. JMIR Ment Health; 2015, 2(1):e8. DOI: 10.2196/mental.3889. PMID: 26543914
13. Fleming T, Bavin L, Lucassen M, Stasiak K, Hopkins S, Merry S. Beyond the Trial: Systematic Review of Real-World Uptake and Engagement With Digital Self-Help Interventions for Depression, Low Mood, or Anxiety. J Med Internet Res; 2018, 20(6):e199. DOI: 10.2196/jmir.9275. PMID: 29875089
14. Torous J, Nicholas J, Larsen ME, et al. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. Evidence-Based Mental Health; 2018, 21:116-119. DOI: 10.1136/eb-2018-102891. PMID: 29871870.
15. Baumeister H, Reichler L, Munzinger M, Lin J. The impact of guidance on Internet-based mental health interventions — a systematic review. Internet Interv; 2014, 1(4):205-215. https://doi.org/10.1016/j.invent.2014.08.003
16. Z Xiao, MX Zhou, QV Liao, G Mark, C Chi, W Chen, H Yang. Tell Me About Yourself: Using an AI-Powered Chatbot to Conduct Conversational Surveys. arXiv preprint arXiv:1905.10700. 2019. https://arxiv.org/pdf/1905.10700.pdf
17. Lucas GM, Rizzo A, Gratch J, et al. Reporting mental health symptoms: breaking down barriers to care with virtual human interviewers. Front Robot AI; 2017, 451:1-9. https://doi.org/10.3389/frobt.2017.00051
18. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining gamification. Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments 2011 ACM; 15th International Academic MindTrek Conference envisioning Future Media Environments; 2011, 9–15
19. Cugelman B. Gamification: what it is and why it matters to digital health behavior change developers. JMIR Serious Games; 2013, 1(1):e3. doi:10.2196/games.3139. PMID: 25658754
20. Moreno AL, DeSousa DA, Souza AMFLP, Manfro GG et al. Factor structure, reliability, and item parameters of the brazilian-portuguese version of the GAD-7 questionnaire. Temas psicol; 2016, 24(1):367-376. http://dx.doi.org/10.9788/TP2016.1-25.
21. Osório FL, Mendes AV, Crippa JAS, Loureiro SR. Study of the discriminative validity of the PHQ-9 and PHQ-2 in a sample of Brazilian women in the context of primary health care. Perspectives in Psychiatric Care; 2009, 45:216-227. https://doi.org/10.1111/j.1744-6163.2009.00224.x
22. Patias ND, Machado WL, Bandeira DR, Dell'Aglio DD. Depression Anxiety and Stress Scale (DASS-21) - Short Form: Adaptação e Validação para Adolescentes Brasileiros. Psico-USF; 2016, 21(3):459-469. https://doi.org/10.1590/1413-82712016210302
23. Castro LS. Adaptation and validation of the Insomnia Severity Index (ISI): population characteristics, normative values and associated factors. Dissertação apresentada a Universidade Federal de São Paulo – Escola Paulista de Medicina – para obtenção do Título de Mestre em Ciências, 2011.

http://repositorio.unifesp.br/bitstream/handle/11600/23193/Tese14242.pdf?sequence=1&isAllowed=y
24. Schuster MS, Dias VV. Oldenburg Burnout Inventory - validação de uma nova forma de mensurar Burnout no Brasil. Ciência & Saúde Coletiva; 2018, 23(2):553-562.
25. Mello AMG, Chor Dóra, Faerstein Eduardo, Lopes Claudia de S, Werneck Guilherme Loureiro. Versão resumida da "job stress scale": adaptação para o português. Revista de Saúde Pública; 204, 38(2):164-171. https://doi.org/10.1590/S0034-89102004000200003
26. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A Brief Measure for Assessing Generalized Anxiety Disorder: The GAD-7. Arch Intern Med; 2006, 166(10):1092–1097. doi:10.1001/archinte.166.10.1092. PMID: 16717171
27. Lovibond PF, Lovibond SH. The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck depression and anxiety inventories. Behaviour Research and Therapy; 1995, 33(3): 335-343. doi: 10.1016/j.rbp.2012.05.003
28. Bastien CH, Vallières A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. Sleep Med; 2001, 2:297–307.
29. Demerouti, E., & Bakker, A. B. (2008). The Oldenburg Burnout Inventory: A good alternative to measure burnout and engagement. In J. Halbesleben (Ed.), Handbook of stress and burnout in health care (pp. 65–78). Hauppauge, NY: Nova Science Publishers.
30. Eysenbach G. (2004). Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). Journal of medical Internet research, 6(3), e34. https://doi.org/10.2196/jmir.6.3.e34.
31. Viana MC, Teixeira MG, Fidel F, Bassani IS, Andrade LH. São Paulo Megacity Mental Health Survey - a population-based epidemiological study of psychiatric morbidity in the São Paulo metropolitan area: aims, design and field implementation. Brazilian Journal of Psychiatry; 2009, 31(4):375-386. https://doi.org/10.1590/S1516-44462009000400016. PMID: 20098829
32. Galea S, Tracy M. Participation rates in epidemiologic studies. Ann Epidemiol; 2007, 17:643–53. DOI: 10.1016/j.annepidem.2007.03.013. PMID: 17553702
33. Hooley T, Wellens J, Madge C, Goss S. Online Research Methods for Mental Health. In: Anthony, K, Mertz Nagel, D and Goss, S (eds.) The Use of Technology In Mental Health: Applications, Ethics and Practice; 2010. Springfield: Charles C Thomas.
34. World health Organization (WHO). Depression and other common mental disorders: Global Health Estimates [Internet].2017 Apr 7 [cited 16 June 2017]. www.who.int/mental_health/management/depression/prevalence_global_health_estimates/en/
35. Silva, M.T., Galvão, T.F., Martins, S.S., Pereira, M.G.. Prevalence of depression morbidity among Brazilian adults: a systematic review and meta-analysis. Brazilian Journal of Psychiatry; 2014, 36(3). DOI: 10.1590/1516-4446-2013-1294. PMID: 25119639
36. Moreira DS, Magnago RF, Sakae TM, Magajewski FRL. Prevalência da síndrome de burnout em trabalhadores de enfermagem de um hospital de grande porte da região

sul do brasil. Cad. Saúde Pública; 2009, 25(7):1559-1568. https://doi.org/10.1590/S0102-311X2009000700014. PMID: 19578577

37. Alves PC, Oliveira AdF, Paro HBMdS. Quality of life and burnout among faculty members: How much does the field of knowledge matter? PLoS ONE; 2019, 14(3):e0214217. https://doi.org/10.1371/journal.pone.0214217. PMID: 30901373
38. Karyotaki, E., Kleiboer, A., Smit, F., Turner, D., Mira, A., Andersson, G., Berger, T., Botella, C., Bretón-López, J., Carlbring, P., Christensen, H., Graaf, E., Griffiths, K., Donker, T., Farrer, L., Huibers, M., Lenndin, J., Mackinnon, A., Meyer, B., Cuijpers, P. Predictors of treatment dropout in self-guided web-based interventions for depression: An 'individual patient data' meta-analysis. Psychological medicine, 2015, 45. DOI: 10.1017/S0033291715000665. PMID: 25881626
39. Smail-Crevier R, Powers G, Noel C, Wang J. Health-Related Internet Usage and Design Feature Preference for E-Mental Health Programs Among Men and Women. *J Med Internet Res*.; 2019, 21(3):e11224. doi:10.2196/11224. PMID: 30882361
40. Becker, S., Miron-Shatz, T., Schumacher, N., Krocza, J., Diamantidis, C., Albrecht, U.V. mHealth 2.0: Experiences, Possibilities, and Perspectives. JMIR Mhealth Uhealth; 2014, 2(2), e24. 10.2196/mhealth.3328. PMID: 25099752.