What depression symptoms matter to stakeholders of ecological momentary assessment studies?

Giba, Ágoston

Citation

Giba, Á. (2022). *What depression symptoms matter to stakeholders of ecological momentary assessment studies?*.

Universiteit Leiden

Psychologie
Faculteit der Sociale Wetenschappen

What depression symptoms matter to stakeholders of ecological momentary assessment studies?

Ágoston Giba

Master Thesis Psychology

Institute of Psychology

Faculty of Social and Behavioural Sciences – Leiden University

Unit: Clinical Psychology MSc

Date: 21/08/2022

Student number: 1819348

Thesis Supervisor: Eiko Freid

Abstract

Depression places an immense burden on our society as a whole, with young people being disproportionately affected. Current depression assessment instruments have remained mostly unchanged in the past decades despite their numerous shortcomings, such as the lack of symptom overlap between scales and recall bias caused by the use of retrospective self-reports. Ecological momentary assessment (EMA), an alternative assessment method provides a solution to recall bias by repeatedly sampling participants in real time. High participant burden is a risk factor that needs to be combated by narrowing down items of assessment to the ones that are most important for diagnosing depression. The aim of this study was to utilize stakeholders to narrow down depression symptoms to the ones most relevant to people and to investigate the cause for these symptoms' high personal relevance. A quantitative analysis integrating the frequency of symptoms selected as personally relevant and their ranking of relevance found symptoms Pleasure Loss, Fatigue, Pessimism, Sad Mood, and Anxiety to be the most relevant symptoms. A subsequent qualitative analysis of these five symptoms showed that a symptom's level of disability and severity in daily life, as well as their role as a barrier to recovery, are named as the main cause for high personal relevance. Furthermore, stakeholders named a variety of causal relations between symptoms that challenge the current static view of depression symptoms. Our study provided a proof-of-principle that stakeholders are a viable and valuable data source for EMA depression studies.

Table of Contents

1 What depression symptoms matter to stakeholders of ecological momentary assessment studies?

Major depressive disorder (MDD) is defined as a pervasive state of sad mood and/or the lack of ability to experience pleasure and interest, lasting 2 weeks or longer, accompanied by a wide range of somatic & cognitive changes that impair the individual's ability to function (American Psychiatric Association, 2013; Malhi & Mann, 2018). The word depression is used to indicate a range of phenomena in literature, in this thesis, it is used to encompass a broader sense of groups of depressive symptoms that are also represented in depressive disorders outside of MDD, as well as by clinical, subclinical or residual symptoms. Depression is markedly different from natural mood fluctuations and short-lived emotional responses to challenges that we experience in our everyday life. Factors that distinguish depression from these are the severity, extent, and manifestation of symptoms (Moreh & O'Lawrence, 2016).

Below, we first discuss the burden depression places on our society, followed by a brief introduction of the way depression is currently assessed and details on its limitations, as well as their repercussions. Next, a new source of data, stakeholder engagement, and an emerging assessment method, ecological momentary assessment (EMA), are introduced as solutions for the aforementioned limitations of current depression assessment methods. We conclude by describing the current study's aim to unify these emerging disciplines.

1.1 Depression Prevalence

An estimated 4% of the global population, approximately 322 million people suffer from depression worldwide (World Health Organization, 2017). Its lifetime prevalence is 15-18%, with 40% of people experiencing their first episode of depression before the age of 20, with a median of 25 years, and each episode greatly increases the chance of further episodes occurring (Malhi & Mann, 2018). Research on the prevalence of depression over time has been divisive in the past decade, with some reporting an increasing trend, while other studies show stable prevalence trends with no relevant changes over time (Patten et al., 2015; Weinberger et al., 2018). Yet, recent comprehensive reviews of findings indicate a predominantly increasing trend in the prevalence of depression within virtually all populations, with incidence cases increasing by 49.86% since 1990 (Liu et al., 2020; Moreno-Agostino et al., 2021).

1.2 Depression Etiology

Depression is a condition without a single cause to point to, yet, combinations of specified risk factors have been shown to increase the probability of acquiring the disorder (Moreh & O'Lawrence, 2016). Depression has a marked genetic component, the occurrence of the condition is significantly higher in relatives of depressed people than what can be observed in the general population (Mullins & Lewis, 2017). Aside from genetics; social and cultural factors such as social support, employment, number of children & loss of a family member, as well as socioeconomic status were identified to have a role in the occurrence of depression. Although, it is important to note that the 12-month estimated prevalence of major depressive disorder in high-income countries (5.5%) is comparable with low-income & middle-income countries (5.9%), implying that its occurrence cannot be ascertained as a plain consequence of poverty or living standards alone (Liu et al., 2020). Common unhealthy lifestyle choices, such as smoking & alcohol consumption also appear to heighten the risk of depression (Pavkovic et al., 2018). Pre-existing physical conditions, such as diabetes, cancer, or AIDS are also considered risk factors (Sotelo et al., 2014).

1.3 Consequences of Depression

Depression is ranked by the World Health Organization (WHO) as the largest contributor to disability worldwide, being responsible for 7.5% of all years lost to disability, more than any other mental condition in the world (Liu et al., 2020). Moreover, depression has been identified as the main determinant of suicide deaths, further elevating its global disease burden (Vigo et al., 2016). According to a meta-analysis on suicide attempts, the lifetime prevalence of attempts among people diagnosed with major depressive disorder is 35%, 5-fold higher than in the general population (Dong et al., 2019). The total economic burden of Major Depressive Disorder (MDD) is estimated to be 326.2 billion dollars per year in the United States alone (Greenberg et al., 2021). Depression is associated with high personal and societal costs throughout the life span (Rodríguez et al., 2012). Even mild levels of depression have a major impact on people's quality of life. In depressed outpatients, high indirect costs are also associated with depression due to the resulting productivity loss & unemployment (Wang et al., 2017). These substantial economic costs are a consequence of low rates of treatment & recovery among depressed individuals (Thornicroft et al., 2017).

In spite of the immense strain depression puts on affected individuals & our society, only a minority of people with MDD actually receive any treatment. This treatment gap is even more concerning considering that less than 50% of people with depression recognize that they have the condition and that it is treatable (Thornicroft et al., 2017). This lack of awareness strongly suggests an absence of proper warning systems in place for depression that would clearly communicate to the individual the nature & severity of their condition.

1.4 Effect of Depression on Young Population

Despite the increasing trend in depression prevalence worldwide, young adults seem to be disproportionately affected (Nguyen et al., 2019). Professionals acknowledge the age between 18 and 29 as a critical developmental period where the building of a stable life structure takes place (Gustavson et al., 2018). Depression in early adulthood can lead to an accumulation of negative consequences throughout the entirety of adult life, impacting social relationships, career prospects, and education attainment (Aalto-Setälä et al., 2001 as cited in Ibrahim et al., 2013). Depression is known to increase the risk of suicide in adolescents and young adults, which the WHO recently estimated to be the fourth leading cause of death for 15-29-year-olds worldwide (Moreh & O'Lawrence, 2016; World Health Organization, 2021). For instance, among college students, a reportedly high-risk group for psychological problems, the prevalence of depression reaches as high as 30.6% (Ibrahim et al., 2013). Moreover, recent external events, such as the COVID-19 worldwide pandemic had a major impact on the lives of young adults specifically. A recent large-scale meta-analysis on the effects of the pandemic on mental health found that young people and college students experienced the highest negative psychological impact from the pandemic, including a larger increase in depression rates than any other affected populations (Dragioti et al., 2022).

1.5 Current Measures of Depression

Depression is primarily measured by the use of two main classificatory diagnostic systems; the Diagnostic and Statistical Manual of Mental Disorders (DSM) and the International Classification of Diseases (ICD), DSM being the predominant classificatory system in research & practice (Malhi & Mann, 2018). Both systems rely on the identification of a number of key symptoms under a specified time window. In the most recent iteration, DSM-5, an individual must present five or more specified symptoms out of nine every day for a period of two weeks (American Psychiatric Association & American Psychiatric Association, 2013). Remarkedly, none of the symptoms classifying depression are exclusive to the disorder, also featuring in numerous other psychiatric illnesses. The DSM identifies sad mood & anhedonia as the core symptoms of MDD, requiring at least one of them to be present for diagnosis. The rest of the symptoms include weight loss/gain, insomnia/hypersomnia, psychomotor agitation/retardation, fatigue, worthlessness/guilt, concentration problems/indecisiveness, and suicidal ideation/attempt. These symptoms are given equal significance by the DSM-5. The resulting lack of an established hierarchy among the non-core symptoms is a limitation that may dilute the accurate conceptualization of depression and make its detection more difficult, especially considering that MDD symptoms are shared by numerous other disorders as well (Malhi & Mann, 2018). The integrity of the DSM criteria has been called into question before, which remained mostly unchanged for the past 4 decades, as some symptoms seem to feature in it for more historical reasons than evidence-based ones (Fried et al., 2016).

1.6 Limitations of Screening Instruments

To help clinicians better identify and determine the severity of depression, screening scales have been developed for various clinical settings. However, these measures come with their own set of limitations. The most glaring issue is the high variance in constructs measured across screening scales. A study by Fried (2017) compared seven of the most commonly used depression scales worldwide and found very low overlap between measured symptoms. This high variability in measurement constructs can gravely limit the establishment of a more standardized measurement of depression and results in very different clinical presentations being classified as having depression with seemingly equivalent severity (Malhi & Mann, 2018). When scales measure different constructs, different results are likely, which may significantly bias the results and can make generalizability across scales problematic. Consequently, the long-standing silent assumption that different screening instruments can be used interchangeably to measure depression severity does not seem to hold true. Moreover, this shines a light on the uneasy truth that there is a substantial lack of consensus regarding which constructs are most important to the measurement of depression (Fried et al., 2016). This is a serious issue that severely undermines accurate clinical diagnosis and threatens not only the generalizability but also the validity of large volumes of research conducted on depression under this premise. Furthermore, it creates uncertainty about what the content of depression screening ought to be. Immediate action is required to find a solution.

There are historical reasons for this. Depression itself, throughout its history, has been conceptualized in many different forms; a physical disease, a severe form of grief, a maladaptive psychological response, a set of attitudes, and a myriad of others (Haroz et al., 2017). This precipitates the idea that these vast differences in conceptualization had a major influence on the development of depression screening instruments, thus they mainly reflect the clinical opinion of their time & their creators, therefore potentially biasing the content of the scales toward their conception of depression. For example, a commonly used depression screening instrument, the Beck Depression Inventory (BDI) is based on Beck's cognitive triad theory of depression (Beck & Beamesderfer, 1974), therefore it reflects the higher emphasis on the cognitive aspect of depression compared to other scales.

1.7 Stakeholders

In light of the large differences in their conceptualization & purpose, the lack of construct overlap between depression screening instruments is less surprising, however, its implications remain severe. In order to work toward a list of constructs that are universally important for depression, research needs to focus on increasing the involvement of people that are affected by it. In research, these people are referred to as stakeholders, defined as *"an individual or group who is either responsible for or affected by health- and healthcare-related decisions that can be informed by research evidence"*(Ray & Miller, 2017).

Companies have been utilizing stakeholders in order to improve & tailor their products, as well as to assess opinions on their legitimacy & survival successfully for nearly a century (Grafé-Buckens & Hinton, 1998). Despite its overarching success in the business field, this trend did not spread to the research community, where the use of stakeholder involvement is limited (Ray & Miller, 2017). The National Institute of Mental Health has outlined the importance of involving stakeholders in all aspects of the research pipeline to increase the effectiveness of mental health interventions (Soyster & Fisher, 2019). A review of results by the INVOLVE research collective found that involving the public positively impacts the quality of research (Staley & INVOLVE, 2009). Benefits of stakeholder engagement (SE) in research include increased relevance & feasibility, improved study designs, more effective implementation & dissemination of research findings, and higher influence on future research (Goodman et al., 2020; South et al., 2016). SE elicits a patient-centered research design that benefits from diverse perspectives outside of the traditional research team, leading to findings that are more sensitive to the needs of target groups (Ray & Miller, 2017).

Mental health research could greatly benefit from stakeholder engagement in numerous ways. Directly incorporating their lived experience and first-hand expertise on a given condition may highlight and help overcome particular challenges in mental healthcare (Madden et al., 2021). In depression research, the challenge of incongruence between which aspects of depression clinicians and screening instruments find relevant and which aspects patients consider relevant have been highlighted by numerous studies (IsHak et al., 2011; Madden et al., 2021; Rosenblat et al., 2019). This relates to the aforementioned lack of consensus on which constructs are the most important for depression measurement, which would be an ideal candidate for the use of SE for the following reasons: stakeholders could provide perspective on which constructs have higher validity in clinical practice and inform assessment of depression by highlighting the constructs that depressed individuals consider relevant.

1.8 Ecological Momentary Assessment

A further limitation of mainline depression screening instruments is their reliance on retrospective self-reports. Reports of this kind can readily become victims of recall bias, a common systematic error that reveals the unreliability of our autobiographical memory (Shiffman et al., 2008). Even the process of memory retrieval itself can be subject to bias, for example, studies illustrated that individuals experiencing negative mood, recall significantly more negative information than positive (Kihlstrom et al. 2000 as cited in Shiffman et al., 2008). Indeed a large body of empirical research points to recall being inaccurate and systematically biased. Consequently, this limits our efforts to accurately recognize and understand behavior as it happens in the real world (Shiffman et al., 2008). In the study of depression and its related phenomena, this means that dynamic mechanisms requiring in-vivo assessment, such as the diurnal variation in mood, as well as changes in context-dependent factors cannot be measured accurately with retrospective self-reports, thus failing to assess characteristics of the individual's depression which may be vital in forming an accurate diagnosis and informing treatment (Malhi & Mann, 2018). In response to the lack of real-time assessment in research, alternative report methods are starting to emerge in research, one of which is ecological momentary assessment (EMA).

EMA is an assessment method that is unique in its approach, it specializes in capturing in-vivo experience by assessing individuals' behavior in real-time, as they unfold during their daily lives (Soyster & Fisher, 2019). This is done by repeated sampling of the participant's experience in its natural context throughout various extents of time with varying intensity of assessment. This type of measurement frequently yields an abundant amount of observation data for each participant. By obtaining a large number of data per participant, it is possible to assess dynamic, within-person processes as they unfold over time and across situations, which would not be possible with routine retrospective report methods (Soyster & Fisher, 2019). By doing so, it enhances the ecological validity of findings and effectively eliminates recall bias, allowing for greater generalization of results to participants' real life (Shiffman et al., 2008).

The data obtained through EMA can be either quantitative (e.g., numerical) or qualitative (e.g., free-response). The method of data collection is either active (e.g., written self-report or answering in an app) or passive through worn sensors or mobile phones (Shiffman et al., 2008). In EMA studies, three distinct sampling paradigms are being used; event-contingent (participants answer in response to a predefined event), signal-contingent (participants are randomly prompted to answer in order to get a representative sample of experience), and time-contingent (participants are prompted according to a fixed time schedule) (Soyster & Fisher, 2019). Sampling frequency, similar to the period of sampling, varies extensively, ranging from as many as 60 times per day to once per day, while the sampling period may range anywhere from days to years (Koval & Kuppens, 2012). With today's technology allowing people to use their phones as assessment devices through applications anywhere and anytime, EMA sampling is readily available and cost-feffective (Husky et al., 2014).

1.9 EMA in Depression Research

In depression research, EMA can be utilized in order to measure how symptoms and within-subject changes vary over time. As of today, only a limited number of studies have examined the utility of EMA in assessing depression specifically, however, existing research suggests that the data acquired from EMA is more accurate than from standard self-report instruments. A study by Ben-Zeev et al. (2009) compared the accuracy of retrospective affective reports to affective EMA reports and found that both the depressed and non-clinical groups reported exaggerated positive and negative affect in the retrospective reports compared to EMA reports. In addition to studying affect, EMA has been found to measure activity levels throughout the day more accurately than retrospective measures (Armey et al., 2015). These findings are important as both affect and activity levels are cornerstone symptoms of MDD. Furthermore, literature suggests that EMA is more advantageous in the measurement of numerous behavioral variables associated with depression-related phenomena such as suicide & self-injury. For instance, Thompson et al. (2014) were able to predict suicidal ideation in participants with inter-episode bipolar disorder significantly better by utilizing EMA to measure affect over an eight-week period than via being assessed by experienced clinicians. Important to note that repeated evaluations of suicidality in EMA procedures did not lead to increased levels of suicidal ideation or behavior, implying the lack of any iatrogenic effect even among high-risk suicidal populations (Husky et al., 2014). The utility of EMA for the measurement of depression symptoms has been illustrated by Targum et al., 2021, who used EMA to track changes in depression symptoms in clinical trials. The study found EMA an efficient and valid measure of depression symptoms over time (Targum et al., 2021).

1.10 EMA Limitations

Seemingly, EMA is equipped to overcome the limitations of current depression measurements by providing a more representative sample of the participant’s experience. However, EMA suffers from notable limitations as well. Firstly, there is a concerning lack of clear guidelines or protocol on how to best construct studies that feature EMA, thus there are significant differences in its utilization throughout studies (Soyster & Fisher, 2019). Second, since the EMA process requires consistent repeated compliance from subjects, prolonged assessments can cause an exceptionally high burden for the participant. This so-called assessment burden is a methodological risk factor that may cause a decrease in compliance rates & non-committal answering. Compliance rates have been shown to erode significantly in EMA studies following the second week of data collection (Broderick et al., 2003) Furthermore, assessment-burden was shown to be less tolerable for people with affective disorders, such as depressed individuals (van Genugten et al., 2020).

1.11 Current Study

It is imperative for EMA researchers to try to ease assessment-burden as much as possible in order to ensure high participant compliance rates & valid results. There are numerous ways to reduce this burden. The first and most obvious way is to reduce the number of items included in the assessment. This, however, carries the risk of an incomplete assessment of depression if too many key depression-related constructs are omitted. The second is to improve the wording of items in a way that is most easily understood by participants while still describing their symptoms in a meaningful way. In order to select that handful of depression-related constructs that are the most relevant for depression assessment, we propose involving stakeholders. In this thesis, we utilize stakeholders to inform item selection for EMA studies that assess changes in depression for extended spans of time. By doing so, we aim to provide a proof of principle that stakeholder engagement can serve as a particularly valuable tool in depression research.

2 Methods

2.1 Participants

The study was conducted online using Microsoft Teams video-conference software. Participants (N=10) were college students (Mean age = 20.55, SD = 2.45, range = 18-26), recruited through Leiden University’s research recruiting system, SONA. Across the 10 participants, 9 identified as female and 1 as male. As the study is centered around young adults, a minimum age of 18 and a maximum age of 35 were defined. Participants were drawn from two specific groups of people. Firstly, five participants were drawn from a group of people who were diagnosed with depression, but no longer fulfilled the criteria for a depressive disorder. Second, five participants were drawn from a group of people indirectly affected by depression, as a result of keeping close, long-term contact with an individual diagnosed with depression. This was defined as either a significant other, family member, or close friend with the duration of close contact being ≥ 1 year.

The exclusion criteria were people with an ongoing depressive episode, the BDI-II was utilized as a prescreen instrument in order to avoid any possible iatrogenic effects that sharing detailed accounts of past symptoms could cause. Inclusion criteria comprised 17 points and below on the BDI-II prescreen, as well as direct or indirect experience with depression.

2.2 Procedure

In order to generate a larger list of constructs that are important for depression measurement, we used the 52 depression symptoms across 7 common rating scales identified by Fried (2017). We included 26 symptoms that appeared on at least 3 of the rating scales for the current study. The symptoms were then transformed into corresponding items by the rater with a primary focus on EMA utility; shortness, conciseness, and understandability were key aspects, with an emphasis on making them relatable in most daily situations. An effort was made to try to stay as close to the article's definition of each symptom as possible (ex.: symptom guilt was transformed into the item "I feel guilty"). Specific symptoms that comprised of distinct components were deemed to be better described by more than one item were divided. One is the symptom "anxious", which was divided into 2 specific items to represent the active, specific mental component "worry" and the more passive, non-specific, and primarily physical component "anxious feeling". The other is symptoms "sympathetic arousal" which was similarly divided into a mental component "alert" and a physical one "restless". The resulting items and their corresponding symptoms are presented in Fig. 1. In the remainder of this thesis, the items are referred to by the symptoms they represent. Ethical approval for the study was given by the Leiden University Psychology Research Ethics Committee (*reference num.: 2021-02-05-E.I. Fried-V3-2814*).

Figure 1

Items with corresponding symptoms

Symptom	Item
Fatigue	I feel tired

Sad Mood	I feel sad
Concentration	I have trouble concentrating
Agitation	I feel agitated
Pessimism	I feel discouraged about my future
Appetite Decrease	My appetite is low
Appetite Increase	My appetite is high
Early Insomnia	I have difficulty falling asleep
Middle Insomnia	I have difficulty staying asleep
Late Insomnia	I wake up earlier than I need to
Hypersomnia	I feel drowsy
Pleasure Loss	I don't get pleasure out of the things I usually enjoy
Guilt	I feel guilty
Interest Loss	I lost my interest in other people or things
Libido	I lost interest in sex
Suicidal Ideation	I have thoughts of suicide
Indecisiveness	I have problems making decisions
Weight decrease	I lost weight
Sympathetic Arousal (Physical)	I feel restless
Worthlessness	I feel that I am worthless
Crying	I cry
Sympathetic Arousal (Mental)	I feel alert
Retardation	I feel restricted in thought and movement
Irritable	I feel easily frustrated
Anxiety (Specific)	I worry
Anxiety (Non-specific)	I feel anxious
Gastrointestinal	I have digestive issues

Note.

Initially, potential participants who signed up via SONA filled out the BDI-II questionnaire through Qualitrics as part of the prescreen procedure. Previous studies have found that BDI is the most representative scale among the most commonly used ones and has the largest average overlap with other scales, therefore we elected to use it for our study as well (Fried, 2017). The cut-off point on the scale was determined at 17 points, just below what is interpreted as “borderline clinical depression” by the BDI. To avoid unintentionally biasing applicants’ scores toward a certain direction and to reduce the negative effects of labeling applicants as possibly depressed, it was stated in the information letter that only individuals with scores inside a “predefined threshold” will be invited to participate. An email was sent out to applicants with a score of ≥ 17, explaining that their score fell outside of the range of inclusion.

Following the prescreening procedure, participants were invited for a video-call interview on Microsoft Teams. The call lasted approx. 60 minutes. During the first 15 minutes of the call, participants were asked to fill out a Qualtrics online self-report questionnaire independently, which included rating each of the 26 symptom items based on how they were feeling at the moment on a scale of 0-100. Afterward, the participants were presented with the same list of items, but this time they were asked to select which of these items they found personally relevant to their experience with depression (from this point on, we refer to this item as “personally relevant”), not what they thought would be most relevant to people suffering from depression in general. They were also notified that the following interview will be based on the items they select. There was no limit to how many items participants could select, however, they were informed on the goal of the study was to narrow these items down into a smaller list of items. This was done to allow participants to base their answers on their experiences, whether they experienced a wide range of symptoms or not.

The second part of the interview consisted of a 30-45 minutes long semi-structured interview based on the items that the participant selected as personally relevant. The length of the interview was dependent on the number of items the participant selected, as well as on the length and depth of their disclosure. The interviews were audio-recorded in order to minimize interruptions in the conversation by continuous note-taking & eliminate inaccurate data collection. At the start of the interview, the participant was told a disclaimer on the current state of depression research, the purpose of the interview, and their role in it. Furthermore, the participant was informed of the procedures in place to maintain their anonymity. These included the removal of any identifying information from the recording, the distortion of the participant’s voice to an unrecognizable level, the secure location of the audio which only researchers involved in this study were allowed to access, and the eventual destruction of the audio after analysis, consistent with Leiden University data policy.

During the administration of the interview, all participants were encouraged to speak their minds & share their experiences with depression. Initially, questions about the items’ legibility & comprehensibility were asked to confirm that all of them were understood as intended. Next, the participants were asked about their experiences with depression. Afterward, they were shown their list of selected items and were asked to give their reasoning behind choosing each item and the nature of personal relevance they ascribe to each of them. Next, the participants were asked about their thoughts and feelings regarding each item in order to obtain information on the emotional aspect the participant assigns to them. Furthermore, they were asked to point out which item or items they considered least relevant to their experience. This was done in order to obtain further information on which items could prove negligible to stakeholders. Additionally, the participants were asked if they would change anything in these items (ex.: wording, content), whether there was any item missing in the questionnaire that they would have liked to see included, and their reasoning behind including those items. Before concluding the interview, the participants were asked to rank the top five most important items from the ones they have previously selected in order of importance and their reasoning behind each ranking. At the conclusion of the interview, the participants were thanked for their effort and offered the choice to be redirected to a mental health professional (school counselor, external psychologist) if they wished to further work on their experiences.

2.3 Data Analysis

The two central sources of quantitative data obtained during the experiment were the frequency of items selected as personally relevant and the top five ranking of the selected items. In addition, the frequency of items picked as least relevant was also considered. As the number of items participants would select as least relevant was expected to vary significantly, participants were not asked to rank them.

2.3.1 Quantitative Analysis

In order to quantify the rankings of the top five most important items across participants, a scoring system was created as presented in Fig. 2.

Figure 2

System of scoring for ranked and non-ranked items

Rank	Score Given

Selected (Not ranked)	4
5th place	5
4th place	5.5
3rd place	6
2nd place	6.5
1st place	7

The scoring system gives primary weight to an item being selected, giving it a score of 4. The increment between ranks is 0.5 (5th place is given a score of 5, while 1st place is worth a score of 7). In order to add distance between items being selected as personally relevant without ranking and items featuring in the top five rankings, the lowest rank (5th place) is scored two increments higher than an item that is only “Selected”.

On the other end of the spectrum, the frequency of items selected as least relevant was also measured in an identical manner to the frequency of most relevant items.

2.3.2 Qualitative Analysis

The five highest scoring symptoms based on our scoring system were chosen to be included in the qualitative analysis across all participants to investigate in greater depth the reason why these symptoms are deemed as most personally relevant by stakeholders. Firstly, the author became familiar with the data by conducting and transcribing the relevant parts of the interviews. These transcripts were analyzed by thematic analysis in order to find the most common themes among participants on each symptom and to identify the cause and level of relevance for each. The analysis was conducted in the qualitative data analysis software named Quirkos. Considering that personal relevance can be defined in many ways, data-driven coding was used to let the qualitative data shape the resulting themes and extra attention was paid to data potentially relating to personal relevance, such as severity and disability. Therefore, the coding scheme was based on the answers of the participants. The thematic analysis was primarily used to reveal the common points of personal relevance between participants concerning each specific symptom. The analysis was further aimed at finding common points of conceptualization of each featured symptom. In addition, perceived connections between items were considered. The data was coded under themes relating to depression symptoms, levels of severity, disabilities & factors that influence or are influenced by symptoms. Conceptually similar subthemes were grouped together and assigned a theme that represents the range of contexts (ex.: “hopelessness”, “helplessness”, and “frustration” about oneself were all grouped under “negative self-image”). There were no groups that contained only a single symptom. These themes were used to register the number of mentions and synthesize the data inside a theme. Due to the lack of secondary raters, the resulting descriptions arising from these themes were an attempt to match the wording and conceptualization of the data provided by the participants as closely as possible in order to reduce interpreter bias.

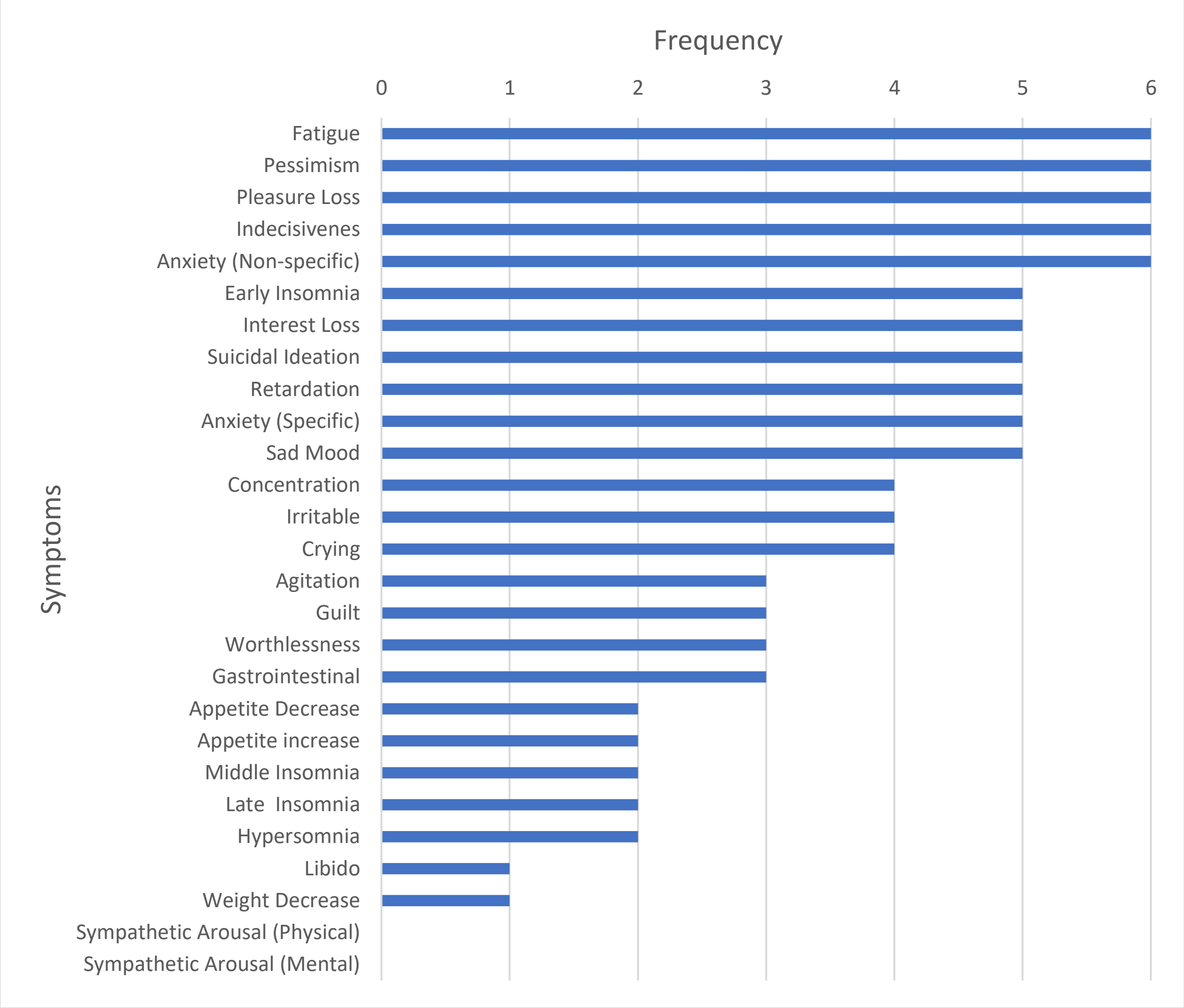
3 Results

3.1 Quantitative Analysis

The frequency of items selected as personally relevant for their experience with depression is shown in Fig. 3. The most frequently selected items (selected 6 times) were: Fatigue, Pessimism, Pleasure Loss, Indecisiveness, and Anxiety (non-specific).

Figure 3

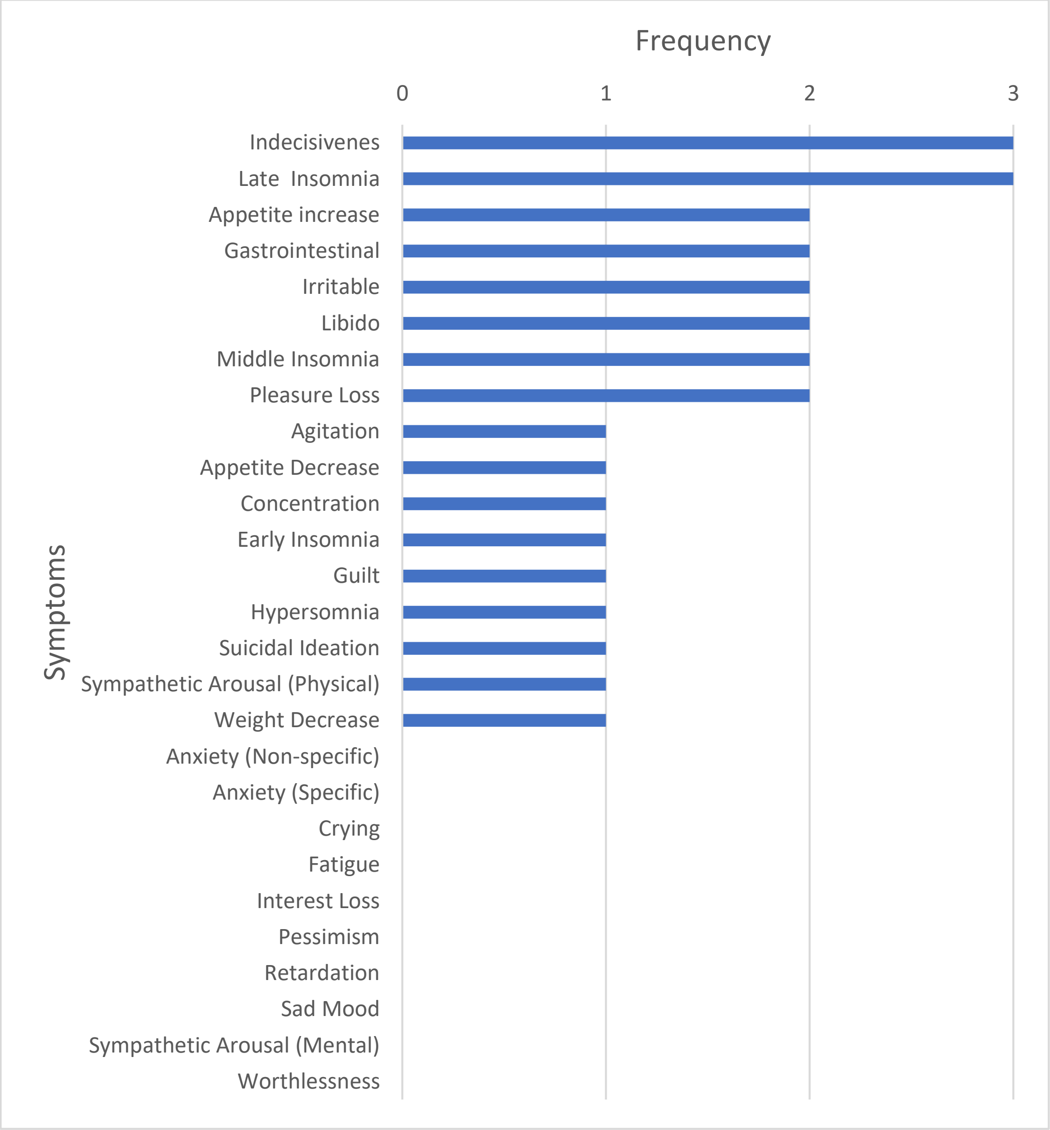
Frequency of symptoms selected as personally relevant



When participants were asked which are the least relevant items for their experience with depression, the most frequently selected (selected 3 times) were: Late Insomnia and Indecisiveness. Due to the freedom given in how many to select, there was high variance in responses; some selected 2 or more symptoms, while others selected none as presented in Fig. 4.

Figure 4

Frequency of symptoms selected as least relevant

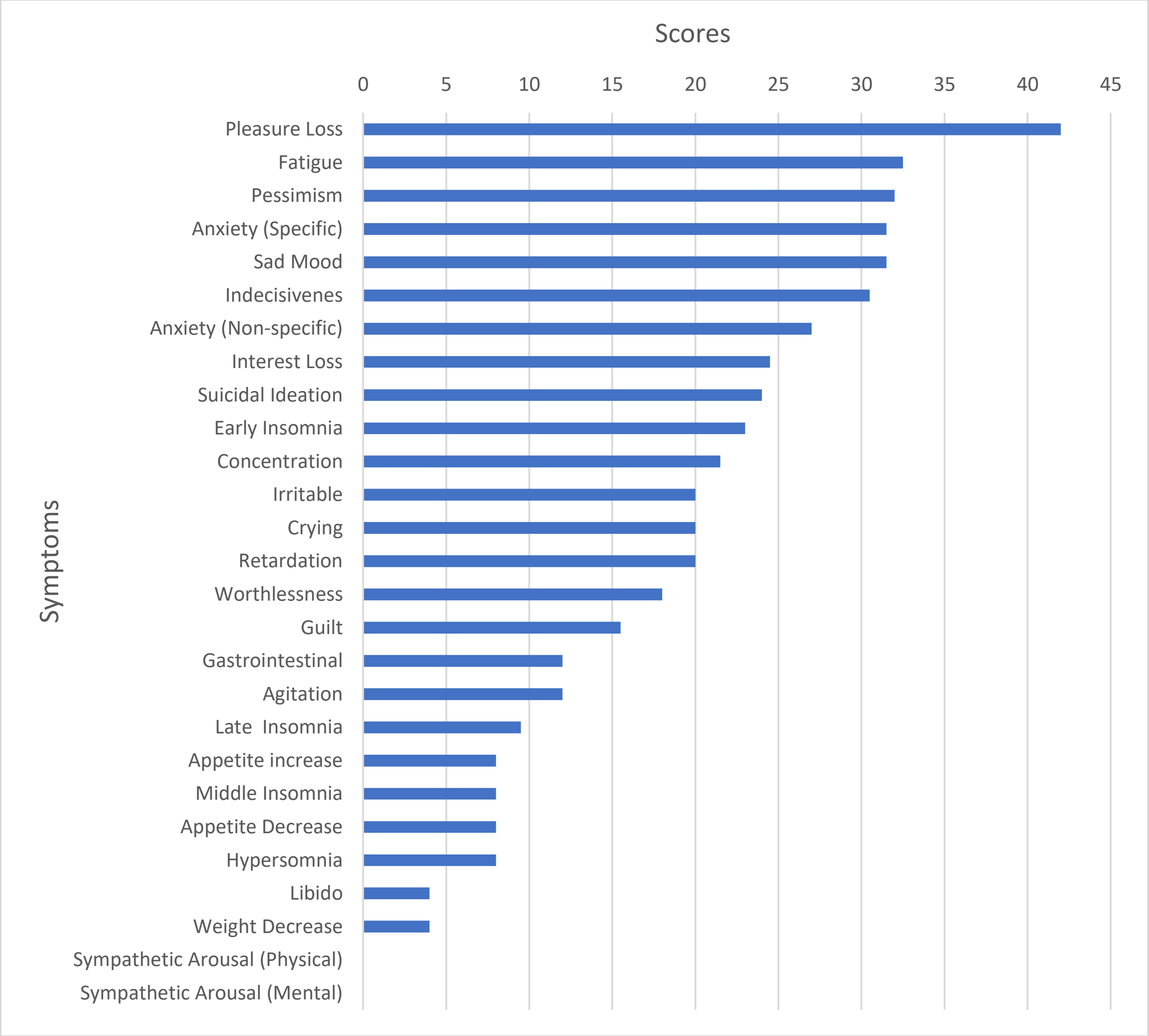


Due to the variability in answering, no clear patterns emerged from symptoms selected as least relevant.

The scoring system described in the Methods aimed to assess the top five rankings whilst also incorporating the frequencies presented in Fig 3. The following five items were selected as most relevant according to our scoring system, in descending order: Pleasure Loss (42 points), Fatigue (32.5 points), Pessimism (32 points), Sad Mood (31.5 points), Anxiety (specific) (31.5 points)

Figure 5

Scoring based on the frequency of selections and top 5 rankings (scoring system)



3.2 Qualitative Analysis

3.2.1 Characteristics of Measurement

Nine out of ten participants were satisfied with the wording & legibility of the items, generally praising their simplicity and coherence. The generic wording of the items was commended by participants, as they saw it as easy to identify with. However, points for improvement were shared by several participants. Firstly, two participants reported items “I feel agitated” and “I feel easily frustrated” as too similar and had difficulty discerning them from each other. Second, the item “I feel restricted in thought & movement” was reported as unclear for two participants, causing confusion. The 0-100 slider implemented in the survey was praised by three participants because of the higher level of nuance it allowed in answering, although, one participant criticized its use as it gave her choice paralysis as a result of having too many options to choose from.

In regards to what items they would have liked to see included but did not, participants reported the need to include items about substance use, irregular sleep patterns, isolation/feelings of being isolated, and physical tension/pain.

3.2.2 Five Most Personally Relevant Symptoms

In the following, we discuss the qualitative investigations of the five most personally relevant symptoms according to our scoring system.

The first symptom, Pleasure Loss, received the highest score based on both selection frequency and top five rankings. Participants described this symptom as numbness or flatline of emotions despite emotion-eliciting circumstances. It was frequently reported by four participants as the largest barrier to recovery, as well as the main cause of contrast between their depressed and non-depressed states. Three participants reported it acting as an alarm for an incoming depressive episode or as predictive of the severity of an incoming episode. Desperation to feel any emotion, including even negative ones was outlined by two participants. Moreover, four participants named it an important mechanism of action in the sequence of depression. These participants reported this symptom to be the elicitor of a vicious cycle of depression symptoms by producing a lack of positive reinforcers in their environment, it elicited hopelessness and desperation and led to experiential avoidance that further reinforced their feelings of hopelessness.

The second highest scoring symptom was Fatigue, which was described as the most prevalent symptom throughout two participants' experiences of depression. For three participants, this symptom was the single most disabling aspect of depression and the starter of the vicious cycle of depression symptoms. Furthermore, two participants reported levels of tiredness to correlate positively with their subjective depressed state, stating that perceived tiredness levels were amplified by sad mood. Three participants reported the symptom to be closely tied to, and in some cases result in, irregular sleep patterns.

The third highest scoring symptom was Sad Mood, which participants described as an all-consuming feeling and a physical weight that accompanies other depression symptoms. Despite this, three participants described feelings of relief and acceptance towards experiencing sadness and voiced a strong preference for it over experiencing Pleasure Loss. One participant described it as acting as a wake-up call for taking action against depression. However, perceiving sadness in depressed, closely-affiliated others was named as the main source of distress for participants indirectly affected by depression.

The fourth highest scoring symptom was Pessimism. This symptom was described by participants as a “tunnel vision” in mainly future perception that is predominantly restricted to potentially negative outcomes. It has been also described as a feeling of hopelessness and generalization of their current depressed state and ability to the future. Participants named numerous exacerbating factors such as uncertainty, high internal & external expectations, comparing oneself to others, and lack of self-efficacy. Furthermore, it was described as being closely interlinked with worrying and resulting in worthlessness by four participants.

The fifth highest scoring symptom was Anxious (specific), described by two participants as a thought loop about personal problems that never end in a solution, one participant referred to it as a vicious cycle of rumination. One participant named it the single most important symptom, that is central to their experience of depression. They stated that other depression symptoms amplify worrying, thus becoming a prevalent symptom that is often experienced as overwhelming and constant, leading to frustration and hopelessness. Sources of worry differed between participants, however, the two most prevalent were future & social perceptions. It has been also described as acting as a signal for an ongoing depressive episode by one participant. Two participants that defined themselves as having a neurotic personality or had pre-existing anxiety problems have noticed a substantial increase in problem-related thoughts and worrying during a depressive episode.

4 Discussion

Despite the vast amounts of empirical research done on the assessment and treatment of depression, advancements have been limited (Fried et al., 2022). The traditional method of depression assessment still employed by today’s most widely used depression instruments; retrospective self-report, suffers from pervasive systematic errors, such as recall bias, thus severely limiting its ecological validity (Shiffman et al., 2008). Ecological momentary assessment (EMA) provides a solution for the errors of retrospective self-report by repeatedly sampling participants over time in the real world. To ease the participation burden of this method, this qualitative study utilized stakeholders, an underutilized data source, through one-on-one interviews to narrow down the 26 most frequently overlapping depression-related symptoms into a smaller list for accessible EMA assessment. To our knowledge, this is the first study to utilize stakeholders in the item selection process of depression symptoms for EMA.

4.1 Summary of Findings

From the initial 26 symptoms provided to participants, the five symptoms deemed most personally relevant were Pleasure Loss, Fatigue, Sad Mood, Pessimism, and Anxious (Specific). Participants differed in their reaction to each symptom, however, certain components were echoed in similar regards throughout the interviews. Components that participants named as most important in being relevant to their depression included the level of disability a symptom caused in daily life, its prevalence throughout a depressive episode, its subjective severity, and its perceived effect as a barrier to recovery. The most relevant symptoms were also reported to serve an alarm function for either the imminence or severity of the incoming episode or serve as a call to action to initiate treatment. Upon looking at the results and analyzing the subgroups of directly and indirectly affected people with depression, no dramatic differences emerged. As such, we aim not to overinterpret these two groups given the small sample size of the study. Three findings stood out that require a more detailed discussion: the pronounced role of the symptom Pleasure Loss, the similarity of perceived causal relations to existing depression models, and the high occurrence of causal relations and perceived interactions between the symptoms.

4.2 The Role of Pleasure Loss in Depression

The symptom Pleasure Loss, coupled with Interest Loss (together often referred to as “anhedonia”) is outlined by DSM-5 as one of the core symptoms of MDD. Despite its importance to the diagnosis of depression, the underlying neurobiological and psychological mechanisms of anhedonia remain poorly understood, therefore it is rarely targeted in depression interventions (Alexander et al., 2019; Watson et al., 2020). Neurobiologically, a large issue that slows down progress in the development of new treatments targeting anhedonia is that the brain mechanisms which elicit the phenomena are still largely unknown

(Alexander et al., 2019). Current literature describes anhedonia as acting upon three types of reward processing: reward wanting, reward learning, and reward liking, the latter being comparable to Pleasure Loss (Borsini et al., 2020). In our study, Pleasure Loss and Interest Loss were decoupled and measured separately, showing that stakeholders placed considerably higher personal relevance on Pleasure Loss. According to stakeholder accounts, the main reason behind this is the perceived centrality of the symptom in the causal cycle of their depression. Pleasure Loss was seen as the initiator of subsequent depression symptoms, as well as the central agent in the maintenance of depression. Perceived severity of the symptom among stakeholders was therefore high, some even voiced the acceptance and marked preference for Sad Mood, the other core symptom of MDD in DSM-5, instead of having to experience Pleasure Loss. The high level of personal relevance and perceived severity stakeholders attribute to Pleasure Loss, coupled with the strikingly modest amount of literature on the role of anhedonia in depression and the limited progress in available literature all imply that Pleasure Loss or reward liking, and anhedonia as a whole, requires significantly more scientific attention to better understand its role in depression onset and maintenance. Furthermore, neuropsychological models that describe the mechanism of action by which each subtype of reward processing is influenced by anhedonia are necessary for the future to inform depression treatment in a meaningful way.

4.3 Causal relations as depression models

All participants described some sort of causal relationship between at least three of the five items they ranked as personally relevant. Many described a vicious cycle elicited by fatigue or emotional numbness acting as a demotivator, leading to avoidance. This avoidance evokes increased worrying and results in a negative future perspective that causes feelings of sad mood, hopelessness & worthlessness. Consequently, these feelings reinforce the lack of motivation. Remarkably, the cycle described by the participants maps appropriately onto two widely recognized psychological models of depression with slight distinctions. Foremost, it shows a very close resemblance to the behavioral activation model of cognitive behavior therapy (CBT) by Lewinsohn (Dimidjian et al., 2011). Two key differences to note, participants more frequently reported Sad Mood as a consequence and eventual reinforcer of lack of motivation, rather than a direct source of lowered behavioral activation (avoidance) and some stated their tiredness starting their depressive cycle instead of the lack of reinforcement. Likewise, participants’ accounts closely map onto the broaden & build theory model of positive psychology by Fredrickson (2001). In this model, the tunnel vision that participants described is an active component defined as the narrowing of the thought-action repertoire that interacts with worrying to produce a negative perspective about the future. However, a major distinction from participants’ accounts is that this model posits sadness to be the elicitor of this narrowing effect on the thought-action repertoire, instead of serving as a consequence. This implies a possibility for stakeholders to inform and augment a variety of psychological theories, improving their ecological validity.

In psychometric depression assessment, causal connections and interactions between symptoms, akin to the ones described by our stakeholders, are primarily informed by concurrent prevalence rates, order of occurrence, sequential comorbidity, and treatment outcomes. However, documenting the presence and severity of symptoms does not provide information on their etiological significance, which could provide increased content validity and inform general theory on depression. Furthermore, assessing clients’ own causal attributions about their depression symptoms have been shown valuable in informing several aspects of the treatment process, such as case conceptualization, differential diagnosis, and treatment planning (Belzer & Schneier, 2004; Frewen et al., 2012). In the current conceptualization and diagnosis of depression, these causal connections remain unexplored. Frewen et al. (2012) devised a methodology called perceived causal relations (PCR) scaling that measures the degree to which participants attribute each symptom as the cause of other symptoms, as well as exploring the directionality of these associations. This provides a framework as well as an accessible methodology for assessing the perceived causal relationships stakeholders provide in the future.

4.4 Network Approach to Depression

The high level of reported causal relations and interactions between the top five items indicates that depression symptoms are not only more interconnected than they are currently represented in literature but that they show a high level of interaction with each other. The current nine criterion symptoms in DSM-V for MDD do not account for such connections, likely contributing to the fact that it is currently among the least reliable diagnoses in DSM-5 field trials (Regier et al., 2013).

An emerging methodological approach aims to rectify this, promptly named the "network approach" by Borsboom (2008), by placing the focus on the associations between symptoms, and investigating the degree to which associations between symptoms mediate a diagnosis. His network theory of psychopathology conceptualizes mental disorders as clusters of symptoms that are unified by causal relations (Borsboom, 2017). Therefore, symptoms do not merely indicate depression, they are a causal agent in themselves that influences other symptoms. Thus, both the individual system elements and their relationships need to be measured, which necessitates a move away from fixed, retrospective assessments in favor of more dynamic ones, such as EMA (Fried et al., 2022). Network analysis could even help conceptualize treatment by instead of targeting a symptom, as it is traditionally done, providing opportunities to target the links between symptoms, thereby changing the structure of the network and promoting lasting change (Borsboom et al., 2021). The aforementioned perceived causal relations scaling method by Frewen has been used to test the direction and magnitude of cause-and-effect associations between participant-reported depression, anxiety, and posttraumatic stress symptoms in a multi-symptom network which found strong moderation effects between PCR scores and symptom frequency scores (Frewen et al., 2013). This indicates that perceived causal relations by stakeholders are predictive of the strength of associations between symptoms, therefore carry significant diagnostic value, and the network approach can be successfully utilized by methods such as PCR to investigate these associations.

4.5 Limitations

The present study aimed to provide a demonstration of the feasibility of utilizing stakeholders in depression research, specifically in the use of EMA studies, based on a limited number of interviews. Thus, the study has several important limitations. First, the sample is small and not representative of people directly and indirectly affected by depression: nine of 10 participants were female, therefore the results may only represent depression constructs that are relevant to female stakeholders. Furthermore, the sample comprised of solely college students therefore it may not be representative of the young adult population, only the young college population. Future stakeholder research is needed to be done with a more diverse population that represents a genuine sample of the young adult population. In addition, the item "I feel discouraged about my future" may only be a highly relevant item in the college student population. Second, the study only utilized one coder, therefore the danger of interpretation bias is evident. Third, the exclusive use of stakeholders in remission as with retrospective and hindsight, they may find different aspects of their depression relevant than stakeholders currently suffering from depression. Fourth, the study did not utilize a screening procedure to screen out comorbid disorders, therefore it is unknown how other disorders influenced the stakeholders' experience of depression and whether depression was only a symptom of another disorder. Fifth, the study did not screen for previous psychotherapy/psychiatric treatment, thus the detailed & introspective accounts participants gave could have been the result of intensive psychoeducation from these treatments. Future research investigating personal accounts should enquire about the extent of the treatment and the information stakeholders received thoroughly. Sixth, due to the autonomy offered to the participants in how many symptoms they wish to select as least personally relevant, no clear patterns emerged from the results. In case future studies opt for investigating this aspect, a structured method of assessment is advised. Finally, the lack of results between groups of stakeholders with direct and indirect experience with depression indicates that the study was too underpowered to show any real differences between these groups. The found differences were highlighted in the results section but given the qualitative nature of the study and the small sample size, replication with larger sample size is necessary to draw meaningful conclusions.

4.6 Conclusion

The present study served as hypothesis generation, with an aim to initiate the first step toward utilizing an often neglected but potentially invaluable data source, stakeholder engagement, for depression research and to demonstrate its potential value. Despite the study's exploratory nature, stakeholders not only provided valuable information regarding which depression symptoms are most relevant to their experience with depression and why, they also supplied information about causal relationships between these symptoms. Such information could be profoundly valuable in informing testing methodology, diagnostic theory, and treatment processes. Hence, our study provided a proof-of-principle that stakeholder engagement can be an exceedingly valuable and versatile tool in depression research. Naturally, the results of this study will need to be confirmed by a larger follow-up study in the future. Depression research is experiencing a surge in utilizing emerging methodologies, such as EMA and the network approach, which would benefit greatly from integrating the use of stakeholders as a data source.

References

Alexander, L., Gaskin, P. L. R., Sawiak, S. J., Fryer, T. D., Hong, Y. T., Cockcroft, G. J., Clarke, H. F., & Roberts, A. C. (2019). Fractionating Blunted Reward Processing Characteristic of Anhedonia by Over-Activating Primate Subgenual Anterior Cingulate Cortex. *Neuron*, *101*(2), 307-320.e6. https://doi.org/10.1016/j.neuron.2018.11.021

American Psychiatric Association, & American Psychiatric Association (Eds.). (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed). American Psychiatric Association.

Armey, M. F., Schatten, H. T., Haradhvala, N., & Miller, I. W. (2015). Ecological momentary assessment (EMA) of depression-related phenomena. *Current Opinion in Psychology*, *4*, 21–25. https://doi.org/10.1016/j.copsyc.2015.01.002

Beck, A. T., & Beamesderfer, A. (1974). Assessment of Depression: The Depression Inventory. *Psychological Measurements in Psychopharmacology*, *7*, 151–169. https://doi.org/10.1159/000395074

Belzer, K., & Schneier, F. R. (2004). Comorbidity of Anxiety and Depressive Disorders: Issues in Conceptualization, Assessment, and Treatment: *Journal of Psychiatric Practice*, *10*(5), 296–306. https://doi.org/10.1097/00131746-200409000-00003

Ben-Zeev, D., Young, M. A., & Madsen, J. W. (2009). Retrospective recall of affect in clinically depressed individuals and controls. *Cognition & Emotion*, *23*(5), 1021–1040. https://doi.org/10.1080/02699930802607937

Borsboom, D. (2008). Psychometric perspectives on diagnostic systems. *Journal of Clinical Psychology*, *64*(9), 1089–1108. https://doi.org/10.1002/jclp.20503

Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, *16*(1), 5–13. https://doi.org/10.1002/wps.20375

Borsboom, D., Deserno, M. K., Rhemtulla, M., Epskamp, S., Fried, E. I., McNally, R. J., Robinaugh, D. J., Perugini, M., Dalege, J., Costantini, G., Isvoranu, A.-M., Wysocki, A. C., van Borkulo, C. D., van Bork, R., & Waldorp, L. J. (2021). Network analysis of multivariate data in psychological science. *Nature Reviews Methods Primers*, *1*(1), 58. https://doi.org/10.1038/s43586-021-00055-w

Borsini, A., Wallis, A. S. J., Zunszain, P., Pariante, C. M., & Kempton, M. J. (2020). Characterizing anhedonia: A systematic review of neuroimaging across the subtypes of reward processing deficits in depression. *Cognitive, Affective, & Behavioral Neuroscience*, *20*(4), 816–841. https://doi.org/10.3758/s13415-020-00804-6

Broderick, J. E., Schwartz, J. E., Shiffman, S., Hufford, M. R., & Stone, A. A. (2003). Signaling does not adequately improve diary compliance. *Annals of Behavioral Medicine*, *26*(2), 139–148. https://doi.org/10.1207/S15324796ABM2602_06

Dimidjian, S., Barrera, M., Martell, C., Muñoz, R. F., & Lewinsohn, P. M. (2011). The Origins and Current Status of Behavioral Activation Treatments for Depression. *Annual Review of Clinical Psychology*, *7*(1), 1–38. https://doi.org/10.1146/annurev-clinpsy-032210-104535

Dong, M., Zeng, L.-N., Lu, L., Li, X.-H., Ungvari, G. S., Ng, C. H., Chow, I. H. I., Zhang, L., Zhou, Y., & Xiang, Y.-T. (2019). Prevalence of suicide attempt in individuals with major depressive disorder: A meta-analysis of observational surveys. *Psychological Medicine*, *49*(10), 1691–1704. https://doi.org/10.1017/S0033291718002301

Dragioti, E., Li, H., Tsitsas, G., Lee, K. H., Choi, J., Kim, J., Choi, Y. J., Tsamakis, K., Estradé, A., Agorastos, A., Vancampfort, D., Tsiptsios, D., Thompson, T., Mosina, A., Vakadaris, G., Fusar-Poli, P., Carvalho, A. F., Correll, C. U., Han, Y. J., … Solmi, M. (2022). A large-scale meta-analytic atlas of mental health problems prevalence during the COVID-19 early pandemic. *Journal of Medical Virology, 94*(5), 1935–1949. https://doi.org/10.1002/jmv.27549

Fredrickson, B. L. (2001). The Role of Positive Emotions in Positive Psychology. *American Psychologist*, 9.

Frewen, P. A., Allen, S. L., Lanius, R. A., & Neufeld, R. W. J. (2012). Perceived Causal Relations: Novel Methodology for Assessing Client Attributions About Causal Associations Between Variables Including Symptoms and Functional Impairment. *Assessment*, *19*(4), 480–493. https://doi.org/10.1177/1073191111418297

Frewen, P. A., Schmittmann, V. D., Bringmann, L. F., & Borsboom, D. (2013). Perceived causal relations between anxiety, posttraumatic stress and depression: Extension to moderation, mediation, and network analysis. *European Journal of Psychotraumatology*, *4*(1), 20656. https://doi.org/10.3402/ejpt.v4i0.20656

Fried, E. I. (2017). The 52 symptoms of major depression: Lack of content overlap among seven common depression scales. *Journal of Affective Disorders, 208*, 191–197. https://doi.org/10.1016/j.jad.2016.10.019

Fried, E. I., Epskamp, S., Nesse, R. M., Tuerlinckx, F., & Borsboom, D. (2016). What are “good” depression symptoms? Comparing the centrality of DSM and non-DSM symptoms of depression in a network analysis. *Journal of Affective Disorders*, *189*, 314–320. https://doi.org/10.1016/j.jad.2015.09.005

Fried, E. I., Flake, J. K., & Robinaugh, D. J. (2022). Revisiting the theoretical and methodological foundations of depression measurement. *Nature Reviews Psychology*, *1*(6), 358–368. https://doi.org/10.1038/s44159-022-00050-2

Goodman, M. S., Ackermann, N., Bowen, D. J., & Thompson, V. S. (2020). Reaching Consensus on Principles of Stakeholder Engagement in Research. *Progress in Community Health Partnerships : Research, Education, and Action*, *14*(1), 117–127. https://doi.org/10.1353/cpr.2020.0014

Grafé-Buckens, A., & Hinton, A.-F. (1998). Engaging the stakeholders: Corporate views and current trends. *Business Strategy and the Environment*, *7*(3), 124–133. https://doi.org/10.1002/(SICI)1099-0836(199807)7:3<124::AID-BSE148>3.0.CO;2-R

Greenberg, P. E., Fournier, A.-A., Sisitsky, T., Simes, M., Berman, R., Koenigsberg, S. H., & Kessler, R. C. (2021). The Economic Burden of Adults with Major Depressive Disorder in the United States (2010 and 2018). *PharmacoEconomics*, *39*(6), 653–665. https://doi.org/10.1007/s40273-021-01019-4

Gustavson, K., Knudsen, A. K., Nesvåg, R., Knudsen, G. P., Vollset, S. E., & Reichborn-Kjennerud, T. (2018). Prevalence and stability of mental disorders among young adults: Findings from a longitudinal study. *BMC Psychiatry*, *18*(1), 65. https://doi.org/10.1186/s12888-018-1647-5

Haroz, E. E., Ritchey, M., Bass, J. K., Kohrt, B. A., Augustinavicius, J., Michalopoulos, L., Burkey, M. D., & Bolton, P. (2017). How is depression experienced around the world? A systematic review of qualitative literature. *Social Science & Medicine*, *183*, 151–162. https://doi.org/10.1016/j.socscimed.2016.12.030

Husky, M., Olié, E., Guillaume, S., Genty, C., Swendsen, J., & Courtet, P. (2014). Feasibility and validity of ecological momentary assessment in the investigation of suicide risk. *Psychiatry Research*, *220*(1–2), 564–570. https://doi.org/10.1016/j.psychres.2014.08.019

Ibrahim, A. K., Kelly, S. J., Adams, C. E., & Glazebrook, C. (2013). A systematic review of studies of depression prevalence in university students. *Journal of Psychiatric Research*, *47*(3), 391–400. https://doi.org/10.1016/j.jpsychires.2012.11.015

IsHak, W. W., Greenberg, J. M., Balayan, K., Kapitanski, N., Jeffrey, J., Fathy, H., Fakhry, H., & Rapaport, M. H. (2011). Quality of Life: The Ultimate Outcome Measure of Interventions in Major Depressive Disorder. *Harvard Review of Psychiatry*, *19*(5), 229–239. https://doi.org/10.3109/10673229.2011.614099

Koval, P., & Kuppens, P. (2012). Changing emotion dynamics: Individual differences in the effect of anticipatory social stress on emotional inertia. *Emotion*, *12*(2), 256–267. https://doi.org/10.1037/a0024756

Liu, Q., He, H., Yang, J., Feng, X., Zhao, F., & Lyu, J. (2020). Changes in the global burden of depression from 1990 to 2017: Findings from the Global Burden of Disease study. *Journal of Psychiatric Research*, *126*, 134–140. https://doi.org/10.1016/j.jpsychires.2019.08.002

Madden, J. M., Foxworth, P. M., Ross-Degnan, D., Allen, K. G., Busch, A. B., Callahan, M. X., Lu, C. Y., & Wharam, J. F. (2021). Integrating Stakeholder Engagement With Claims-Based Research on Health Insurance Design and Bipolar Disorder. *Psychiatric Services*, *72*(2), 186–194. https://doi.org/10.1176/appi.ps.202000177

Malhi, G. S., & Mann, J. J. (2018). Depression. *The Lancet*, *392*(10161), 2299–2312. https://doi.org/10.1016/S0140-6736(18)31948-2

Moreh, S., & O’Lawrence, H. (2016). Common risk factors associated with adolescent and young adult depression. *Journal of Health and Human Services Administration*, *39*(2), 283–310. https://www.jstor.org/stable/44504679

Moreno-Agostino, D., Wu, Y.-T., Daskalopoulou, C., Hasan, M. T., Huisman, M., & Prina, M. (2021). Global trends in the prevalence and incidence of depression:a systematic review and meta-analysis. *Journal of Affective Disorders*, *281*, 235–243. https://doi.org/10.1016/j.jad.2020.12.035

Mullins, N., & Lewis, C. M. (2017). Genetics of Depression: Progress at Last. *Current Psychiatry Reports*, *19*(8), 43. https://doi.org/10.1007/s11920-017-0803-9

Nguyen, C., Murray, G., Anderson, S., Filipowicz, A., & Ingram, K. K. (2019). In vivo molecular chronotyping, circadian misalignment, and high rates of depression in young adults. *Journal of Affective Disorders*, *250*, 425–431. https://doi.org/10.1016/j.jad.2019.03.050

Patten, S. B., Williams, J. V. A., Lavorato, D. H., Fiest, K. M., Bulloch, A. G. M., & Wang, J. (2015). The Prevalence of Major Depression is Not Changing. *The Canadian Journal of Psychiatry*, *60*(1), 31–34. https://doi.org/10.1177/070674371506000107

Pavkovic, B., Zaric, M., Markovic, M., Klacar, M., Huljic, A., & Caricic, A. (2018). Double screening for dual disorder, alcoholism and depression. *Psychiatry Research*, *270*, 483–489. https://doi.org/10.1016/j.psychres.2018.10.013

Ray, K. N., & Miller, E. (2017). Strengthening stakeholder-engaged research and research on stakeholder engagement. *Journal of Comparative Effectiveness Research*, *6*(4), 375–389. https://doi.org/10.2217/cer-2016-0096

Regier, D. A., Narrow, W. E., Clarke, D. E., Kraemer, H. C., Kuramoto, S. J., Kuhl, E. A., & Kupfer, D. J. (2013). DSM-5 Field Trials in the United States and Canada, Part II: Test-Retest Reliability of Selected Categorical Diagnoses. *American Journal of Psychiatry*, *170*(1), 59–70. https://doi.org/10.1176/appi.ajp.2012.12070999

Rodríguez, M. R., Nuevo, R., Chatterji, S., & Ayuso-Mateos, J. L. (2012). Definitions and factors associated with subthreshold depressive conditions: A systematic review. *BMC Psychiatry*, *12*(1), 181. https://doi.org/10.1186/1471-244X-12-181

Rosenblat, J. D., Simon, G. E., Sachs, G. S., Deetz, I., Doederlein, A., DePeralta, D., Dean, M. M., & McIntyre, R. S. (2019). Treatment effectiveness and tolerability outcomes that are most important to individuals with bipolar and unipolar depression. *Journal of Affective Disorders*, *243*, 116–120. https://doi.org/10.1016/j.jad.2018.09.027

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological Momentary Assessment. *Annual Review of Clinical Psychology*, *4*(1), 1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

Sotelo, J. L., Musselman, D., & Nemeroff, C. (2014). The biology of depression in cancer and the relationship between depression and cancer progression. *International Review of Psychiatry*, *26*(1), 16–30. https://doi.org/10.3109/09540261.2013.875891

South, A., Hanley, B., Gafos, M., Cromarty, B., Stephens, R., Sturgeon, K., Scott, K., Cragg, W. J., Tweed, C. D., Teera, J., & Vale, C. L. (2016). Models and impact of patient and public involvement in studies carried out by the Medical Research Council Clinical Trials Unit at University College London: Findings from ten case studies. *Trials*, *17*(1), 376. https://doi.org/10.1186/s13063-016-1488-9

Soyster, P. D., & Fisher, A. J. (2019). Involving stakeholders in the design of ecological momentary assessment research: An example from smoking cessation. *PLOS ONE*, *14*(5), e0217150. https://doi.org/10.1371/journal.pone.0217150

Staley, K., & INVOLVE (Organization), N. I. for H. R. (2009). *Exploring impact: Public involvement in NHS, public health and social care research*. National Institute for Health Research.

Targum, S. D., Sauder, C., Evans, M., Saber, J. N., & Harvey, P. D. (2021). Ecological momentary assessment as a measurement tool in depression trials. *Journal of Psychiatric Research*, *136*, 256–264. https://doi.org/10.1016/j.jpsychires.2021.02.012

Thompson, W. K., Gershon, A., O'Hara, R., Bernert, R. A., & Depp, C. A. (2014). The prediction of study-emergent suicidal ideation in bipolar disorder: A pilot study using ecological momentary assessment data. *Bipolar Disorders*, *16*(7), 669–677. https://doi.org/10.1111/bdi.12218

Thornicroft, G., Chatterji, S., Evans-Lacko, S., Gruber, M., Sampson, N., Aguilar-Gaxiola, S., Al-Hamzawi, A., Alonso, J., Andrade, L., Borges, G., Bruffaerts, R., Bunting, B., de Almeida, J. M. C., Florescu, S., de Girolamo, G., Gureje, O., Haro, J. M., He, Y., Hinkov, H., … Kessler, R. C. (2017). Undertreatment of people with major depressive disorder in 21 countries. *British Journal of Psychiatry*, *210*(2), 119–124. https://doi.org/10.1192/bjp.bp.116.188078

van Genugten, C. R., Schuurmans, J., Lamers, F., Riese, H., Penninx, B. W. J. H., Schoevers, R. A., Riper, H. M., & Smit, J. H. (2020). Experienced Burden of and Adherence to Smartphone-Based Ecological Momentary Assessment in Persons with Affective Disorders. *Journal of Clinical Medicine*, *9*(2), 322. https://doi.org/10.3390/jcm9020322

Vigo, D., Thornicroft, G., & Atun, R. (2016). Estimating the true global burden of mental illness. *The Lancet Psychiatry*, *3*(2), 171–178. https://doi.org/10.1016/S2215-0366(15)00505-2

Wang, J., Wu, X., Lai, W., Long, E., Zhang, X., Li, W., Zhu, Y., Chen, C., Zhong, X., Liu, Z., Wang, D., & Lin, H. (2017). Prevalence of depression and depressive symptoms among outpatients: A systematic review and meta-analysis. *BMJ Open*, *7*(8), e017173. https://doi.org/10.1136/bmjopen-2017-017173

Watson, R., Harvey, K., McCabe, C., & Reynolds, S. (2020). Understanding anhedonia: A qualitative study exploring loss of interest and pleasure in adolescent depression. *European Child & Adolescent Psychiatry*, *29*(4), 489–499. https://doi.org/10.1007/s00787-019-01364-y

Weinberger, A. H., Gbedemah, M., Martinez, A. M., Nash, D., Galea, S., & Goodwin, R. D. (2018). Trends in depression prevalence in the USA from 2005 to 2015: Widening disparities in vulnerable groups. *Psychological Medicine*, *48*(8), 1308–1315. https://doi.org/10.1017/S0033291717002781

World Health Organization. (2017). *Depression and Other Common Mental Disorders.* https://apps.who.int/iris/bitstream/handle/10665/254610/WHOMSD?sequence=1

World Health Organization. (2021). *Suicide*. https://www.who.int/news-room/fact-sheets/detail/suicide