

Common Measures in Mental Health Science: Generalised Anxiety Disorder-7 (GAD-7)

Background

In June 2020, the National Institute of Mental Health and Wellcome [reached a landmark agreement](#) to require the use of a common set of measurement tools in the research they fund. Since then, the International Alliance of Mental Health Research Funders has supported a wide group of funders and journals to join the effort, supported by an expert advisory board.

The [fundamental mission](#) of the Common Measures in Mental Health Science initiative is to ensure that research leads to tangible improvements in the lives of people who experience mental health issues. Given the current, fragmented landscape of mental health data, there is a need to take pragmatic action to make mental health research easier to compare, communicate and interpret.

In the interest of transparency and collaboration, we are sharing the anonymised notes from the discussions that the Common Measures in Mental Health Science Advisory Committee (CMA) had on the GAD-7 on October 27, 2021.

The CMA were asked to share

- a. Their views on the concepts that underlie the questions in the GAD-7
- b. Their views on how well these concepts are captured by the questions
- c. Their experience of using these questions with different populations
- d. What they thought might be missing and needed to get core concepts

Contents

The overall performance of GAD-7 across contexts.....	2
Views on the concepts that underlie GAD-7 questions.....	3
References.....	8

The overall performance of GAD-7 across contexts

The nature of the questions

There were versions of the questions that were in different formats. The ones we selected are all in the same format with the hope that would be helpful to these conversations.

The GAD-7 questions are no more or less objectionable than the PHQ-9 questions.

Issues of context and cultural transferability

The context of the research project and the person responding can shape the responses. A lot of the questions are contextual trying to get at a feeling of recognition. They are difficult to interpret if you don't experience a problem. It is important to be aware of and empirically look at contextual information (e.g. epidemiological vs clinical).

All of the questions can be understood by the populations that researchers are working with, but the issues are not necessarily conceptualised in this way and, thus, they do not always make sense (e.g. the concept of relaxing for rural communities). Hence, researchers, instead of using an idiom, have, at times, found a relevant comparative concept trying to get at the same point. GAD questions don't seem to work well in translation as they require a lot of adaptation.

The reliability of self-assessed measures is questioned in people with entrenched severe mental illness. Some adaptations to circumstances become very normal (e.g. not working, food insecurity) and many questions can be answered normally if clinically relevant psychosis is not assessed. For example, the lives of some people living with severe mental illness in Chennai, India, who are disenfranchised and homeless, are completely at odds with what is identified through this measure's constructs.

However, in a US (or broader Westernised) context, with college student populations, or people living with bipolar or psychosis, the instrument is working fine and seems to correlate with the constructs we expect.

The relationship to help-seeking

The construct of lack of insight in psychosis is common and it gets used in other psychiatric disorders as well. The question of "help-seeking" versus "non-help-seeking" has introduced systematic bias in clinical trials and in clinical treatment settings where there are people who want to be there because they have interpreted their own experiences as problematic, and others who are not interpreting theirs as a problem. Thus, these measures work fine in very particular contexts in which a framing is already set in a particular population. Then, in diverse populations, where many people are "non-health-seeking", there can be issues with the measures.

The relationship to functional impairment

What may be more comparable across contexts is measuring functional impairment (e.g. tenth item of PHQ-9 and WHODAS 2.0).

WHODAS 2.0 can be useful, but it sets a low bar for what counts as functioning. There are many subtle but pervasive impacts of anxiety and depression (e.g. one gets out of the house but feels terrible all the time). The self-identified issue of impact is complicated but important. For example, in primary care settings in Ethiopia, adding the functional impairment question to GAD-7 reduces “prevalence” by more than half (when used with a validated cut-off).

Views on the concepts that underlie GAD-7 questions

Question structure: *Over the last two weeks, how often have you been bothered by the following problems?*

Response scale: *not at all / several days / more than half the days / nearly every day*

1. Feeling nervous, anxious, or on edge

The questions feel fairly straightforward in American English.

It includes three different descriptions. Thinking of translations, it's curious how easily “on edge” can be translated into different languages (e.g. in German, there's not really a term for that), but it's been translated many times so it looks like people have succeeded. Multi-barrelled questions are never good. The concept of “on edge” is an idiom, that doesn't necessarily mean the same thing in different places. Not supportive of any question that includes idioms which are even harder to conceptualize than questions that include emotions.

Idiomatic language is difficult and off-putting when moving across settings. When you have common cultural references, the issue of taking a measure across cultures is that you lose those anchors.

Note:

- The German version of GAD-7 translates this as “nervousness, anxiety, tension”.
- The Amharic version of GAD-7 translates this as “nervous or anxious”.

Situational context would matter a lot e.g. a woman in Afghanistan. In psychosis research, GAD-7 is used more rarely as there is extreme overlap with paranoia and all these questions could be responded to positively. In real-world research, comorbidity is high and there is a risk of measuring something else than what you think you are. This needs to be dealt with in the data harmonisation context.

In some research, the subjects can understand at some level what we're really asking, even though the words that we're using are coarse or uninformative. This is true in anxiety research, but not in other cases. For example, the diagnostic constructs have been confused with research subjects with a history of wartime or refugee trauma.

There is an issue in how questions are used in different studies that are questioning different populations together. The Common Data project may end up starting to look at these types of questions. By collecting this measure with people living with psychosis or previous trauma, we could get insights into how we might make a different set of more precise questions.

GAD-7 is not properly validated in Ethiopia, only been used in a couple of studies. But one thing to note is that anxiety scales have been much easier to translate.

Anxiety is a concept or a construct that people understand a little more readily than depression.

A hypothesis is that the physiological manifestations of anxiety (e.g. adrenaline, fight or flight) are more universal. There's greater difficulty in other feelings or idiomatic expressions. In addition, symptoms of worry, concern, and anxiety are more easily talked about and less stigmatised. It is common to talk about worrying and there is less stigma associated with someone who is worried.

Anxiety may be a clearer concept and a more universal physical experience.

For some immigrants (e.g. German), it can take a while to figure out what people mean when they say "I feel anxious". In this context, some may culturally understand depression better.

"On edge" is an idiom in only some cultures.

Similar issues with the basic expression ("bothered by") as in the discussion about PHQ-9.

There's a question of how differences in language contribute to differential test functioning, especially regarding the impact of bias at scale across groups in different contexts with the same level of underlying depression or anxiety.

Most funders are not allowing researchers to make any edits to questions. They have to collect the question exactly as it's written and deposit the data in that way. They are then free to ask an alternate version of the question in addition to this one.

When running a data archive, having the standard questions along with the additional options is the best way to find out what works in the data.

Option of moving towards having an official translated version and then encouraging future experimentation.

The way that the question is asked does impact how the item is functioning.

2. Not being able to stop or control worrying

If Question 1 is endorsed to some degree, then Question 2 seems to not differ much and be a little more specific.

This question is about agency or control. In the context of psychotic symptoms, this is a critical piece of measurement. It's about having an experience that you can't exercise agency over.

A positive aspect is the addition of the qualifier "more than expected" or "more than usual"; this makes it more specific on why it's "abnormal".

This is an important point when we come to the discussion of "how would you make a better measure".

This is possibly the most critical item in GAD-7 in terms of identifying clinical cases.

3. Worrying too much about different things

May be more difficult to understand what the focus is, especially in contexts of high anxiety.

Everybody might worry too much about some things, but if you're worrying too much about a whole lot of different things that's more suggestive of pathology. A point for concern if experienced researchers are not arriving at the same conclusion.

In a high-adversity setting, worrying about many things could be justified, but still perceived as "too much" by others. Mental health literacy would affect where the line is drawn between "normal" and "abnormal".

In this context, the lack of functional impairment measures is a gap. The concept of adaptive concern: if everybody is just in a constant state of high stress, at a certain point it's just going to start to seem normal or justifiable.

As it's currently written, this question wouldn't necessarily be at all helpful in a population that's not help-seeking. You would ask a different question if you wanted to in that population or perhaps you wouldn't ask this question at all.

There is a broader issue of lack of consensus on what can be seen as "normal" or "abnormal" or what can be defined as disabling (e.g. in different contexts, such as following bereavement, with past child abuse, when living in poverty, or in different regional settings).

Things are messier in multi-diagnostic populations. For example, if someone is having a paranoid thought and is anxious about being followed, then the real psychopathology is the paranoia, and, hence, the anxiety is a completely understandable response.

The measure is relative to the reality of our own world. The focus would be on levels of anxiety that are disabling to someone and they want to have them sorted in different ways. It should not necessarily be about what counts as justifiable or not.

There's some subjectivity about what "too much" is and how people interpret that. There is some ambiguity in the word "different", as in English it can also mean unusual or other multiple things. Especially for patients with underlying OCD.

A potential mismatch between the expectation of the question and the confusion that the word can cause (e.g. worrying about one specific thing, not about many different things).

This question might be valuable in differentiating worry from social anxiety and other types of anxiety.

4. Trouble relaxing

The idea of relaxing is not easy in a setting which does not talk about leisure and it depends on how it gets translated.

This is a more Western/industrialized-country concept.

In some contexts, this has been adapted or conceptually translated to contextual points, such as "doing things that bring you joy" (instead of relaxing where this makes less sense). It is a different concept but related enough.

The question is not really about "joy". It's about being able to calm down and it can be used in this way.

The Spanish translation has worked consistently across language, gender, time, and socioeconomic group.

5. Being so restless that it is hard to sit still

In some studies, there is an initial qualitative process to understand and derive local symptoms of distress, and then compared which questions from common measures best matched the locally described ones. GAD-7 did not perform best in these processes, with few exceptions.

For PHQ-9 and GAD-7 there is a variety of translations to languages. Even for the same language (e.g. French), the translations may be for use in different countries. But as the words are the same, for the common data element of the work, a plan was to add which language the data were collected in.

There is not a standard translation into all languages researchers are working in. In some studies, investigators might want to do the translation a little bit differently, with good reason and room should be left for more than one translation into a single language.

The translations checked are not necessarily idiomatic adaptations but sometimes are random. A small degree of random “noise” may not have a great impact.

6. Becoming easily annoyed or irritable

It could have been phrased as becoming “too” easily annoyed or irritable.

In Ethiopia, this was a good indicator of depression when depression was validated, but it could also be an idiom of anxiety and distress.

Anger should be carefully translated as it could be counter-cultural and might be perceived as critical.

“Bothered by” as a construct is similar to “being irritable”, which in a high-stress context would be frequent. The concept of being bothered by your irritability more than usual is the one that is more indicative of depressive mood.

When “or” is used in a question, it can be used to describe something in a different way versus something else, which may not be clear here, but it does not seem to be a major issue.

7. Feeling afraid, as if something awful might happen

The question in English seems very clear.

This question does not work in areas of ongoing conflict or violence because there are legitimate reasons to be fearful.

The question depends a lot on the baseline of where you're living and what you're experiencing at that moment.

The question has not raised much problem in translation.

Clinically has not caused problems. “Fear” is a very clear concept in an American context and is a critical discriminating symptom.

There are no good self-reported instruments on fear, and this would be an important area of research for connecting biology to cognitive neuroscience to self-reporting. Anger is not a DSM diagnosis but is a code in ICD.

The possibility of any conceptual construct overlap of fear with various positive symptoms and measures of paranoia can be investigated.

Items related to “scared” and “intolerance” are included in the DASH instrument.

References

Batterham PJ, Sunderland M, Carragher N, Calear AL. Development and community-based validation of eight item banks to assess mental health. *Psychiatry Research*. 2016 Sep 30;243:453-62.

Zhang C, Wang T, Zeng P, Zhao M, Zhang G, Zhai S, Meng L, Wang Y, Liu D. Reliability, validity, and measurement invariance of the general anxiety disorder scale among Chinese medical university students. *Frontiers in Psychiatry*. 2021 May 19;12:648755.