

Working Paper Series
**Comparing estimates of
psychological distress using
7-day and 30-day recall
periods: Does it make a
difference?**

Miranda Chilver
Richard Burns
Ferdie Botha
Peter Butterworth

Working Paper No. 19/22
December 2022

Comparing estimates of psychological distress using 7-day and 30-day recall periods: Does it make a difference?*

Miranda Chilver

National Centre for Epidemiology and Population Health,
The Australian National University

Richard Burns

National Centre for Epidemiology and Population Health,
The Australian National University

Ferdi Botha

Melbourne Institute: Applied Economic & Social Research,
The University of Melbourne
ARC Centre of Excellence for Children and Families over the Life Course

Peter Butterworth

National Centre for Epidemiology and Population Health,
The Australian National University
Melbourne Institute: Applied Economic & Social Research,
The University of Melbourne

**Melbourne Institute Working Paper No. 19/22
December 2022**

* This research was supported by the Australian Government through the Australian Research Council's Centre of Excellence for Children and Families over the Life Course (Project ID CE200100025). Corresponding author: Miranda.Chilver@anu.edu.au.

**Melbourne Institute: Applied Economic & Social Research
The University of Melbourne
Victoria 3010 Australia
T +61 3 8344 2100
F +61 3 8344 2111
E melb-inst@unimelb.edu.au
W melbourneinstitute.unimelb.edu.au**

Melbourne Institute: Applied Economic & Social Research working papers are produced for discussion and comment purposes and have not been peer-reviewed. This paper represents the opinions of the author(s) and is not intended to represent the views of Melbourne Institute. Whilst reasonable efforts have been made to ensure accuracy, the author is responsible for any remaining errors and omissions.

Abstract

Self-report measures are widely used in mental health research and may use variable recall periods depending on the purpose of the assessment. A range of studies aiming to monitor changes in mental health over the course of the COVID-19 pandemic opted to shorten recall periods to increase sensitivity to change over time compared to standard, longer recall periods. However, many of these studies lack pre-pandemic data using the same recall period and may rely on pre-existing data using standard recall periods as a reference point for assessing the impact of the pandemic on mental health. The aim of this study was to assess whether comparing scores on the same questionnaire with a different recall period is valid. A nationally representative sample of 327 participants in Australia completed a 7-day and 30-day version of the six-item Kessler Psychological Distress Scale (K6) and a single-item measure of psychological distress (TTPN item) developed for the Taking the Pulse of the Nation survey. Linear mixed models and mixed logistic regression models were used to assess whether altering the recall period systemically changed response patterns within subjects. No substantive recall period effects were found for either the K6 or the TTPN item total scores, although there was a trend towards higher K6 scores when asked about the past 30 days compared to the past 7 days ($b = 0.25$, 95% CI: -0.01, 0.52). This may have been driven by the “feeling nervous” item which had significantly higher scores for the 30-day compared to the 7-day recall period. Neither the K6 nor the TTPN item were significantly affected by the recall period when reduced to a binary variable of likely severe mental illness. The results indicate that altering the recall period of psychological distress measures does not substantively alter the score distribution in the general population of Australian adults. Future research could investigate whether this is due to stability in psychological distress or recall bias.

JEL classification: I18

Keywords: depression, anxiety, recall periods, mental health

Introduction

Self-report measures are ubiquitous in mental health research as they provide convenient subjective insight to a person's current psychological state. Psychological self-report measures ask participants to report the symptoms they have experienced over a specified period of time. These reporting timeframes or recall periods differ between self-report scales, and different recall periods might be used with the same scale depending on the purpose of the assessment. For instance, the Kessler Psychological Distress Scale was designed using a 30-day recall period (Kessler et al., 2002) but was adapted to a 12-month recall period to make it comparable to other scales included in the US National Health Interview Survey (Kessler et al., 2003), and can be shortened to a 7-day recall period (National Comorbidity Survey, 2022). Similarly, the SF-36 health measure has been adapted from the original 30-day recall period to a 7-day recall period (Keller et al., 1997). Longer recall periods are better suited to capturing less frequent or more stable phenomenon while shorter recall periods are better suited to more frequent or less stable phenomenon (Stull et al., 2009). The recall period can also impact prevalence estimates (Freeman et al., 2010). As prevalence refers to the proportion of people in a population with a specific condition during a certain period, a longer recall period is associated with a higher prevalence as the condition is more likely to have been present during a longer timeframe than a shorter one. However, the correspondence between prevalence and recall period may differ for scale scores that ask about the frequency of symptoms rather than the presence or intensity of symptoms (Winkielman et al., 1998).

The COVID-19 pandemic introduced a period of rapid change and possible instability in mental health in the general population. Because short recall periods are better suited to capturing changes in mental health than longer recall periods (Keller et al., 1997), many studies monitoring the mental health impact of the pandemic opted to use scales with shorter

recall periods of one or two weeks (Batterham et al., 2021; Botha et al., 2022; Fisher et al., 2021; Zhang & Liu, 2021). One example of this is the Australian Taking the Pulse of the Nation (TTPN) survey which measured psychological distress over 50 waves during the COVID-19 pandemic using a 7-day recall period instead of the standard 30-day recall period applied in comparable psychological distress measures (Botha et al., 2022). The aim of the current study was to assess whether changing the recall period from 30 days to 7 days systemically alters psychological distress scores, specifically the widely used 6-item Kessler Psychological Distress Scale (K6) and the single psychological distress item applied in the TTPN survey. Like many other surveys implemented to observe the impact of COVID-19 on mental health, the TTPN project lacks pre-pandemic reference data. In this context, it would be useful to know whether pre-existing population-level data using comparable measures but with a different recall period can be used to provide a reference point for pre-pandemic mental health. For example, the Household, Income and Labour Dynamics in Australia (HILDA) Survey has measured the K6 biennially since 2007 and could provide a pre-pandemic reference point of Australian mental health if the 30-day recall period is comparable to the 7-day recall period.

Recall bias. Self-report measures requiring participants to recall past events are vulnerable to inaccuracies and biases in memory recall (Sato & Kawahara, 2011; Thomas & Diener, 1990). Shorter recall periods tend to result in more accurate reports than longer recall periods because they enquire about more recent events that are remembered more easily. Thomas and Diener (1990) compared participants' daily emotional ratings taken over a period of three or six weeks to their retrospective emotional ratings over the full period. They found that, compared to daily reports, participants tended to overestimate the intensity of both positive and negative emotions in retrospective reports and to over-estimate the frequency of their negative emotions relative to positive emotions. A more recent study similarly reported

that participants overestimated their negative emotions in retrospective reports compared to daily reports, though this may be specific to depression and anxiety symptoms (Sato & Kawahara, 2011). Further investigation suggested that the overestimation of negative emotions was partly explained by participants' most intense emotions during the recall period. As the likelihood of more intense emotions is greater over longer recall periods, this could explain why some studies have found that 30-day recall periods result in higher estimates of negative emotions than 7-day recall periods (Batterham et al., 2019).

Question interpretation. Varying the recall period could also affect responses by changing how participants interpret the survey questions (Igou et al., 2002; Winkielman et al., 1998). Given the brevity of self-report items, participants might look to contextual information such as the recall period to decide what type of information is relevant to the question. Participants may assume a question using shorter recall periods is focused on more frequent, less intense experiences, whereas longer recall periods imply the question is asking about less frequent, more intense experiences (Igou et al., 2002). This can result in participants under-reporting frequencies when asked about longer timeframes compared to shorter timeframes as more frequent events may be viewed as irrelevant to the question.

Effects of scale response options. Although the prevalence of mental disorders is directly related to the reporting timeframe, this may not be the case for self-report measures based on symptom frequency instead of incidence, such as the K6. Specifically, the K6 uses a scale ranging from "none of the time" to "all of the time." Events that never occurred in the past 7 days might have occurred in the past 30 days, increasing the minimum scores. However, events that occurred in the past 7 days did not necessarily occur more frequently over the past 30 days, and in some cases may even have been experienced less frequently when averaged over the longer timeframe. This effect should limit the degree to which total

scale scores increase from the 7-day to 30-day recall period, and most of the change would be expected among individuals selecting “none of the time” on the 7-day version.

Given the importance of identifying appropriate pre-COVID mental health measures, understanding the impact of the recall period on estimates of distress and other psychological health outcomes has new prominence. To provide clarity about whether prevalence of psychological distress can be compared between pre-pandemic measures using a 30-day recall period and post-pandemic measures using a 7-day recall period, this study investigates whether changing the recall period on the K6 and TTPN item changes the distribution of responses. Previous reports comparing 30-day and 7-day recall periods on different self-report measures have shown mixed results, with some measures showing no change in the distribution whereas others show a higher central tendency for 30-day relative to 7-day recall periods (Batterham et al., 2019; Keller et al., 1997). However, if any effect of the recall period exists, it was expected that this would be in the direction of higher scores for longer recall periods.

Method

Participants

A general population sample was recruited through the Australian Online Research Unit (ORU) panel during the first two weeks of December 2021. The ORU is an online survey platform that uses both online and offline recruitment methods to build a regionally representative panel from the general public interested in contributing to research. The ORU randomly selects potential participants who meet the eligibility criteria and provides an incentive for their participation. This study sought adults 18 years of age and older currently residing in Australia with an even split between males and females, and approximately equal representation across age groups. This study did not apply exclusion criteria. Based on a power analysis, a sample of at least 250 participants was sought to have 75% power to detect

an effect size of 0.15 on the TTPN single item scale. The study was approved by the ANU Human Research Ethics Committee (Protocol 2021/736).

Measures

Kessler Psychological Distress Scale. The K10 is a widely used measure of psychological distress and is commonly used to identify serious mental illness (Kessler et al., 2002). Six items from the K10 form the shortened K6 scale; these six items were the focus of the current analysis for comparison with the TTPN survey which included the K6 in two waves. These items ask participants how often they felt: nervous, hopeless, restless or fidgety, so sad that nothing could cheer them up, that everything was an effort, and worthless. Response options range from *none of the time* to *all of the time* on a 5-point scale. Participants were asked to consider the past 7 days and the past 30 days in providing their responses. Scores on the K6 range from 6-30, with scores of 19 or higher indicating high levels of psychological distress and likely severe mental illness (National Comorbidity Survey, 2022).

TTPN psychological distress item. A single-item adaptation of the K6 was developed for use in the TTPN survey as an ultra-short measure of psychological distress (Botha et al., 2022). This item asks participants, “*During the past [recall period], about how often did you feel depressed or anxious?*” with the same response options as the K6. Scores on this measure range from 1 to 5, with scores of 4 or 5 indicating high levels of psychological distress and likely severe mental illness.

The ORU questionnaire also included a measure of age group and sex to be included as covariates in the analysis.

Procedure

All participants completed the K6 and TTPN psychological distress item with both the 7-day and 30-day recall period. The order of the K6 and TTPN and recall period order was

counterbalanced across participants. After completing the K6 and TTPN item (with either the 7-day or 30-day recall period), participants completed an unrelated distractor task before completing the K6 and TTPN item again with the alternate recall period.

Statistical analysis. Linear mixed models correcting for statistical non-independence due to repeated measures were used to test whether mean scores on the K6 and TTPN mental distress item differed according to the recall period. Age and gender were included as covariates in the model. Furthermore, to test whether any potential differences between the scale would be perceived as clinically relevant, a mixed logistic regression model was used to assess whether the recall period would influence the probability of falling into the likely severe mental illness category from either measure.

Results

The responding sample consisted of 330 participants, three of whom were excluded due to missing data. The age and gender distribution of the remaining 327 respondents are provided in Table 1. There was approximately equal distribution across age groups, though with fewer participants aged 45-54 and over 65 compared to other age ranges, and approximately equal numbers of men and women. Mean scores on the K6 and TTPN psychological distress scale by recall period are provided in Table 2. Means and standard deviations for individual items of the K6 and TTPN were similar regardless of the recall period. The prevalence of likely severe mental illness was similar for the K6 and TTPN and both scales showed similar numbers of participants falling into the category regardless of the recall period.

Table 1. Participant characteristics

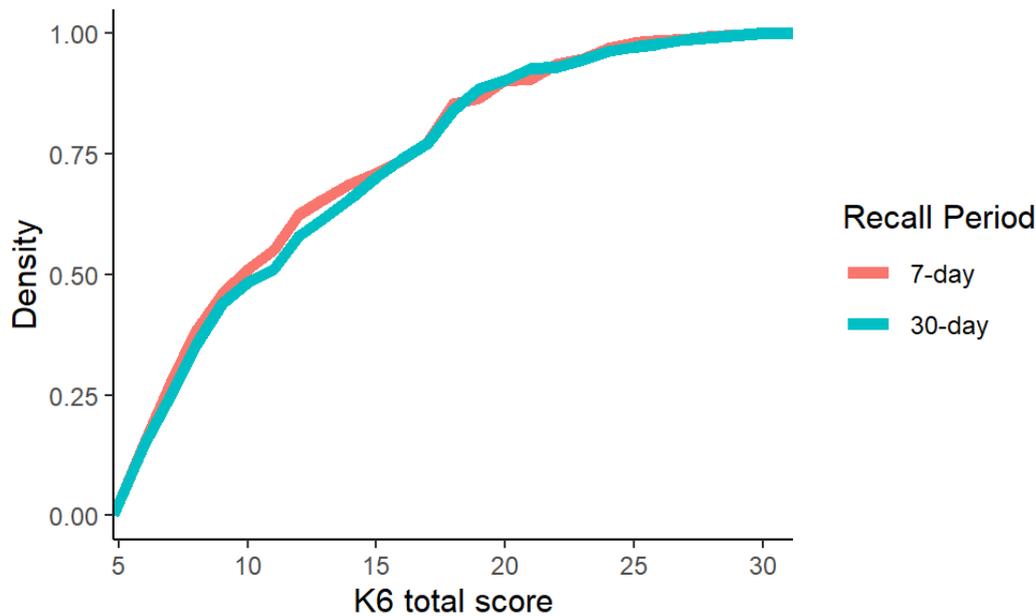
Characteristic	N	%
Age group		
18-24	59	18.0
25-34	57	17.4
35-44	68	20.8
45-54	40	12.2
55-64	63	19.3
65+	40	12.2
Gender		
Male	166	50.8
Female	161	49.2

Table 2. Mean and standard deviation of psychological distress and likely severe mental illness

Scale/item	7-day M (SD)	30-day M (SD)
K6 Total	12.15 (5.79)	12.43 (5.79)
1. Nervous	2.20 (1.07)	2.30 (1.09)
2. Helpless	1.92 (1.10)	1.94 (1.10)
3. Restless or fidgety	2.13 (1.07)	2.16 (1.08)
4. Like everything is an effort	2.21 (1.15)	2.25 (1.14)
5. So sad nothing could cheer up	1.81 (1.12)	1.88 (1.11)
6. Worthless	1.88 (1.11)	1.95 (1.14)
Severe Mental Illness	N = 48 (14.7%)	N = 51 (15.6%)
TTPN Psychological Distress Measure	2.27 (1.11)	2.26 (1.06)
Severe Mental Illness	N = 48 (14.7%)	N = 46 (14.1%)

K6. Linear mixed models showed no substantive effect of the recall period on mean K6 scores, although there was a trend towards higher reported frequencies using the 30-day recall period compared to the 7-day recall period ($b = 0.25, p = 0.059, 95\% \text{ CI: } -0.01, 0.52$). This effect is illustrated in Figure 1. The random effects logistic regression model showed no indication of an effect of the recall period on whether participants fell into the category of likely significant mental illness ($\text{OR} = 1.59, p = 0.408, 95\% \text{ CI: } 0.53, 4.82$). However, consistent with the overall score, there was a slight tendency towards more psychological distress for the 30-day recall period compared to the 7-day recall period.

Figure 1. Cumulative distribution of the K6 total score using a 7-day versus 30-day recall period



Further investigation of each of the individual six items found that mean scores on the item “*how frequently did you feel nervous?*” was significantly but not substantively impacted by the recall period whereby participants tended to respond with higher scores when asked about the past 30 days relative to being asked about the past 7 days ($b = 0.09$, $p = 0.025$, 95% CI: 0.01, 0.17). The cumulative distribution plot of this item, presented in Figure 2, shows that the change is largely driven by a shift from “none of the time” to “a little of the time”. No other items showed a significant effect of the recall period on mean responses.

TTPN mental distress item. Linear mixed models showed no significant effect of recall period on the mean TTPN item score ($b = -0.01$, $p = 0.726$, 95% CI: -0.08, 0.06), and logistic mixed models similarly showed no effect of the recall period on whether participants fell into the category of likely significant mental illness (OR = 0.75, $p = 0.595$, 95% CI: 0.26, 2.15). This was also reflected in the cumulative distribution plot shown in Figure 3, which follows the same pattern in both the 7-day and 30-day recall period.

Figure 2. Cumulative distribution of the K6 “feeling nervous” item using a 7-day versus 30-day recall period

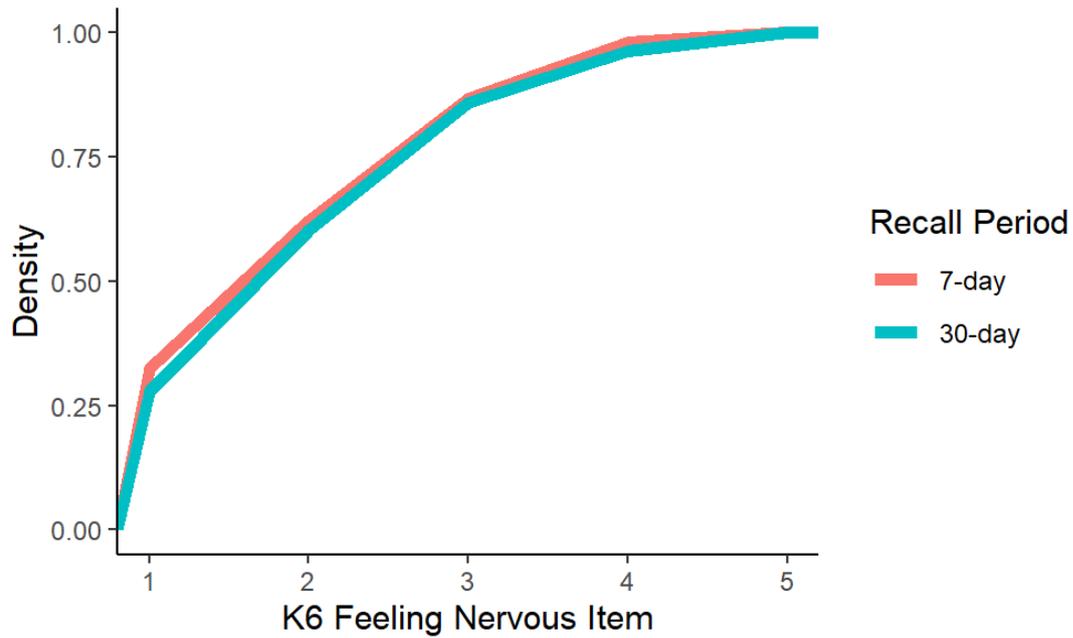
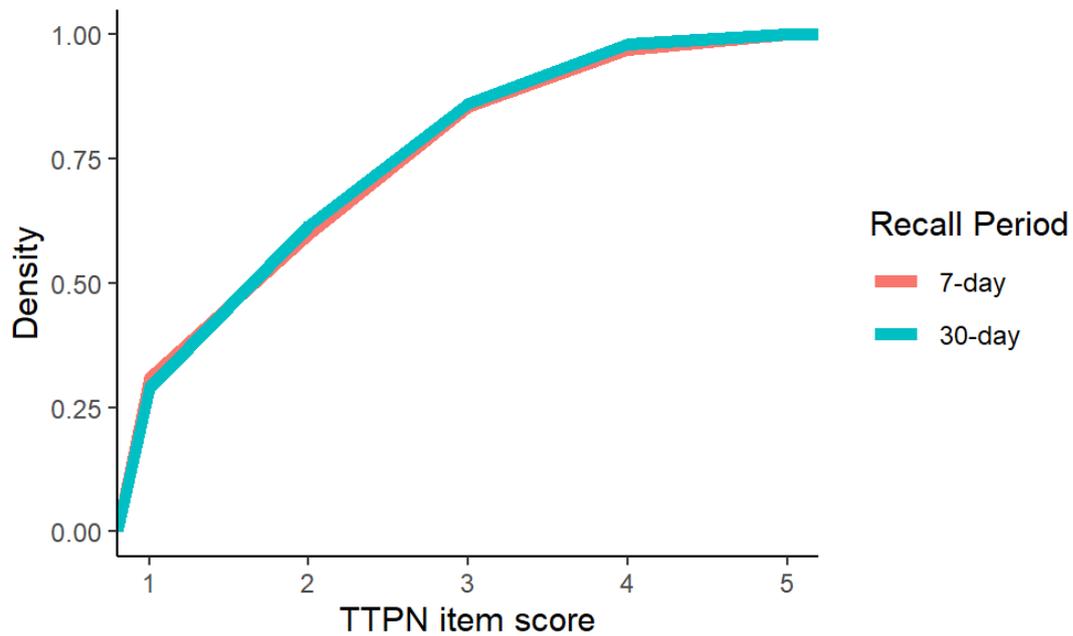


Figure 3. Cumulative distribution of the TTPN psychological distress item using a 7-day versus 30-day recall period



Discussion

The aim of this study was to investigate whether the score distributions of psychological distress according to the K6 and TTPN item are sensitive to whether a 7-day or

30-day recall period is used. The results indicate that any effect of the recall period on mean scores is very small and there is no evidence that the difference between recall periods is of clinical relevance based on no significant change to the binary outcome of likely severe mental illness on either the K6 or the TTPN item. There was a tendency towards higher mean scores when using a 30-day recall period compared to a 7-day recall period, but effect sizes were consistently small and did not reach statistical significance for the total scores. When considering potential practical implications, comparing estimates of psychological distress during the COVID-19 pandemic based on the 7-day recall period with pre-pandemic measures using a 30-day recall period would (if anything) underestimate the mental health impact of the pandemic. However, the scale differences are small enough that the prevalence of likely severe mental illness can be fairly compared between these recall periods for both the K6 and the TTPN item.

Although there was no overall effect of the recall period on mean scores of either the K6 or the TTPN item, one K6 item asking how often respondents felt nervous did show a small but significant increase in the mean score when comparing the 30-day recall period to the 7-day recall period. This is in line with previous studies comparing 30- and 7-day versions of self-report measures that have found trends towards higher scores over longer recall periods (Batterham et al., 2019; Keller et al., 1997). This effect could reflect a tendency to overestimate the occurrence of emotional states over longer recall periods (Sato & Kawahara, 2011; Thomas & Diener, 1990), however, it was not reflected in any other items measuring similar constructs. The cumulative distribution plot shows that the change in the distribution between the 7-day and 30-day version is driven by an upshift of low scorers such that individuals who reported feeling nervous “none of the time” when asked about the past 7 days were more likely to select “a little of the time” when asked about the past 30 days. This item also had one of the highest mean scores overall, suggesting that feeling nervous is

among the more common reported symptoms of psychological distress. As such, it could be that the shift in responses specific to this item simply reflects a higher likelihood of feeling nervous over a month-long period relative to the other symptoms included in the K6. Nevertheless, this shift is not clinically relevant as it predominantly reflects a change in individuals who infrequently experience nervousness.

Future research should investigate why certain emotional states might be more sensitive to changes in the recall period than others. Previous studies using undergraduate samples have suggested that negative emotions are more susceptible to recall bias than positive emotions (Sato & Kawahara, 2011; Thomas & Diener, 1990). However, the current study used a larger, more representative sample and found that only nervousness was significantly affected by the recall period. Another question is whether participants consider the same types of emotional experiences to be relevant to responding to both the 7-day and 30-day versions of the questionnaires. It has been shown previously that obtaining similar mean scores for different recall periods does not necessarily mean that the participants are considering the same type of events in both recall periods (Winkielman et al., 1998). Finally, while the current study sought a representative participant sample, some populations may be missed by the ORU panel. The comparability of the K6 and TTPN psychological distress measures over the 7- and 30-day recall periods might therefore not generalise across all population groups.

In sum, this study found that altering the recall period of the K6 and TTPN psychological distress item from 30 days to 7 days does not substantively alter the score distribution in a general population of Australian adults. An implication of this result is that pre-pandemic measures of psychological distress in Australia using the 30-day recall period can be reasonably compared to post-pandemic measures taken using the 7-day period, pending comparability of samples and research methods. For example, if the HILDA K6

scores using the 30-day recall period were to be used as a reference point of psychological distress in the Australian population, this would provide a conservative comparison to post-pandemic measures taken as part of the TTPN survey using a 7-day recall period. This is especially true when comparing the prevalence of likely severe mental illness as differences in the distribution between the two recall periods was largely at low levels of distress.

However, caution should be taken in interpreting changes in individual items of the K6 scale as some items are more sensitive to the changed recall period than others.

References

- Batterham, P. J., Calear, A. L., McCallum, S. M., Morse, A. R., Banfield, M., Farrer, L., Gulliver, A., Cherbuin, N., Harris, R. M. R., Shou, Y., & Dawel, A. (2021). Trajectories of depression and anxiety symptoms during the COVID-19 pandemic in a representative Australian adult cohort. *The Medical Journal of Australia*, *214*(10), 462–468. <https://doi.org/10.5694/mja2.51043>
- Batterham, P. J., Sunderland, M., Carragher, N., & Calear, A. L. (2019). Psychometric Properties of 7- and 30-Day Versions of the PROMIS Emotional Distress Item Banks in an Australian Adult Sample. *Assessment*, *26*(2), 249–259. <https://doi.org/10.1177/1073191116685809>
- Botha, F., Butterworth, P., & Wilkins, R. (2022). Evaluating how mental health changed in Australia through the COVID-19 pandemic: Findings from the ‘Taking the Pulse of the Nation’ (TTPN) survey. *International Journal of Environmental Research and Public Health*, *19*(1), 558. <https://doi.org/10.3390/ijerph19010558>
- Fisher, J., Tran, T., Hammarberg, K., Nguyen, H., Stocker, R., Rowe, H., Sastri, J., Popplestone, S., & Kirkman, M. (2021). Quantifying the mental health burden of the most severe covid-19 restrictions: A natural experiment. *Journal of Affective Disorders*, *293*, 406–414. <https://doi.org/10.1016/j.jad.2021.06.060>
- Freeman, E. J., Colpe, L. J., Strine, T. W., Dhingra, S., McGuire, L. C., Elam-Evans, L. D., & Perry, G. S. (2010). Public health surveillance for mental health. *Preventing Chronic Disease*, *7*(1), 1–7.
- Igou, E., Bless, H., & Schwarz, N. (2002). Making sense of standardized survey questions: The influence of reference periods and their repetition. *Communication Monographs*, *69*(2), 179–187. <https://doi.org/10.1080/714041712>
- Keller, S. D., Bayliss, M. S., Ware, J. E., Hsu, M. A., Damiano, A. M., & Goss, T. F. (1997). Comparison of responses to SF-36 Health Survey questions with one-week and four-week recall periods. *Health Services Research*, *32*(3), 367–384.
- Kessler, R. C., Andrews, G., Colpe, L. J., Hiripi, E., Mroczek, D. K., Normand, S.-L. T., Walters, E. E., & Zaslavsky, A. M. (2002). Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychological Medicine*, *32*(6), 959–976. <https://doi.org/10.1017/S0033291702006074>
- Kessler, R. C., Barker, P. R., Colpe, L. J., Epstein, J. F., Gfroerer, J. C., Hiripi, E., Howes, M. J., Normand, S.-L. T., Manderscheid, R. W., Walters, E. E., & Zaslavsky, A. M. (2003). Screening for Serious Mental Illness in the General Population. *Archives of General Psychiatry*, *60*(2), 184. <https://doi.org/10.1001/archpsyc.60.2.184>
- National Comorbidity Survey. (2022). *K10 and K6 scales*. https://www.hcp.med.harvard.edu/ncs/k6_scales.php

Sato, H., & Kawahara, J. (2011). Selective bias in retrospective self-reports of negative mood states. *Anxiety, Stress & Coping*, 24(4), 359–367. <https://doi.org/10.1080/10615806.2010.543132>

Stull, D. E., Leidy, N. K., Parasuraman, B., & Chassany, O. (2009). Optimal recall periods for patient-reported outcomes: Challenges and potential solutions. *Current Medical Research and Opinion*, 25(4), 929–942. <https://doi.org/10.1185/03007990902774765>

Thomas, D. L., & Diener, E. (1990). Memory accuracy in the recall of emotions. *Journal of Personality and Social Psychology*, 59(2), 291–297. <https://doi.org/10.1037/0022-3514.59.2.291>

Winkielman, P., Knauper, B., & Schwarz, N. (1998). Looking back at anger: Reference periods change the interpretation of emotion frequency questions. *Journal of Personality and Social Psychology*, 75(3), 719–728. <https://doi.org/10.1037//0022-3514.75.3.719>

Zhang, Y., & Liu, B. (2021). Psychological Distress Among Chinese College Students During the COVID-19 Pandemic: Does Attitude Toward Online Courses Matter? *Frontiers in Psychology*, 12, 665525. <https://doi.org/10.3389/fpsyg.2021.665525>



THE UNIVERSITY OF
MELBOURNE



MELBOURNE
INSTITUTE
APPLIED ECONOMIC &
SOCIAL RESEARCH

60
YEARS
IMPACT