Psychological Assessment

ISSN: 1040-3590

2023, Vol. 35, No. 11, 959–973
https://doi.org/10.1037/pas0001237

Predictors of Time-Varying and Time-Invariant Components of Psychological Distress During COVID-19 in the U.K. Household Longitudinal Study (Understanding Society)

Pascal Schlechter[1], Tamsin J. Ford[1], Sally McManus[2], and Sharon A. S. Neufeld[1]
[1] Department of Psychiatry, University of Cambridge
[2] National Centre for Social Research, University of London

To understand psychological distress during COVID-19, we need to ensure that the same construct is measured over time and investigate how much of the variance in distress is attributable to chronic time-invariant variance compared to transient time-varying variance. We conducted secondary data analyses of Understanding Society, a U.K. probability-based longitudinal study of adults, using prepandemic (2015–2020) and pandemic data (N = 17,761, April 2020–March 2021). Using the General Health Questionnaire–12 (GHQ-12), analyses encompassed (a) five annual waves before COVID-19 plus the first survey wave during COVID-19 and (b) eight (bi)monthly waves during COVID-19. We investigated (a) longitudinal measurement invariance of distress, (b) time-invariant and time-varying variance components of distress using latent trait–occasion modeling, and (c) predictors of these different variance components. In all analyses, unique measurement invariance in distress was established, indicating the same unidimensional construct was measured using the GHQ before and during COVID-19. Time-varying variance was higher at the first COVID-19 lockdown (April 2020, 61.2%) compared to before COVID-19 (~50%), suggesting increased fluctuations in distress at the start of the pandemic. Sensitivity analyses with equal time lags pre- and during COVID-19 confirmed this interpretation. During the pandemic, the highest distress time-varying variance (40.7%) was detected in April 2020, decreasing to 29.0% (July 2020) after restrictions eased. Despite mean-level fluctuations, time-varying variance remained stable during subsequent lockdowns, indicating more rank-order stability after this first major disruption. Loneliness most strongly predicted time-varying variance during the first lockdown. Life dissatisfaction and financial difficulties were associated with both variance components throughout the pandemic.

Public Significance Statement
By demonstrating unique longitudinal measurement invariance before and during the COVID-19 pandemic, this study indicates that the GHQ-12 can be used to reliably measure changes in psychological distress over time, even through periods of significant change to context and environment. Latent trait–occasion modeling suggested increased fluctuations in psychological distress at the start of the pandemic. Loneliness, financial hardship, and low life satisfaction were consistently associated with new onset of psychological distress.

Keywords: COVID-19, psychological distress, latent trait–occasion modeling, longitudinal measurement invariance, General Health Questionnaire

Supplemental materials: https://doi.org/10.1037/pas0001237.supp

Pascal Schlechter https://orcid.org/0000-0002-5916-3694
Pascal Schlechter was funded by the Cusanuswerk. Sharon A. S. Neufeld was supported by the Cundill Centre for Child and Youth Depression and Wellcome Trust Institutional Strategic Support Fund. Tamsin J. Ford received support from the National Institute for Health Research (NIHR) Cambridge Biomedical Research Centre. The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the NIHR, National Health Service, or the Department of Health and Social Care. The authors have no conflicts of interest to disclose.

Pascal Schlechter played a lead role in conceptualization, formal analysis, investigation, methodology, software, visualization, writing–original draft, and writing–review and editing. Tamsin J. Ford played a supporting role in supervision, methodology, and writing–review and editing. Sally McManus played a supporting role in conceptualization and writing–review and editing. Sharon A. S. Neufeld played a lead role in supervision, conceptualization, methodology, writing–original draft, and writing–review and editing.

Correspondence concerning this article should be addressed to Sharon A. S. Neufeld, Department of Psychiatry, University of Cambridge, Douglas House, 18b Trumpington Road, CA CB2 8AH, United Kingdom. Email: sasn2@medschl.cam.ac.uk

Since the onset of the coronavirus disease 2019 (COVID-19) pandemic, relatively high rates of mental health problems have been reported across countries (Xiong et al., 2020). To counteract the rapid spread of COVID-19, governments around the world implemented public health emergency measures such as lockdowns, quarantines, social distancing, and travel restrictions (Sen-Crowe et al., 2020). These measures have been associated with a range of consequences including loneliness, perceived loss of control, financial or employment insecurity, and decreased life satisfaction (Holmes et al., 2020; Magson et al., 2021). Representative cohort studies that compared mental health before the pandemic with mental health during the first lockdowns reported higher mean levels of psychological distress and increased distress above a clinical threshold at the latter point (Patel et al., 2022; Pierce et al., 2020). Specifically, mental distress in the United Kingdom was 8.1% higher in April 2020 compared to the period between 2017 and 2019 (Banks & Xu, 2020). Although mean distress levels decreased in the United Kingdom around July 2020 after COVID-19-related restrictions were eased, distress did not reduce to prepandemic lows (Chandola et al., 2022; Patel et al., 2022). While most U.K. participants had consistently good mental health in the pandemic (Pierce et al., 2021), a substantial proportion experienced elevated distress in a period until May 2021 (Ellwardt & Präg, 2021). These latent trajectory analyses identified participants who were particularly vulnerable, such as those suffering from economic deprivation or living alone (Ellwardt & Präg, 2021; Pierce et al., 2021). Studies have identified further risk and protective factors, such as COVID-19-related worries and feeling socially connected (Magson et al., 2021).

For future prevention and intervention, it is critical to enhance understanding of the antecedents of poor mental health during the COVID-19 pandemic (Holmes et al., 2020). If predictors of change are identified, clinicians and policymakers can be equipped with knowledge about key intervention targets. However, investigating change in psychological distress and identifying risk and protective factors in mental health needs to be grounded in (a) statistically sound models that accurately measure the same construct over time even when circumstances change (Liu et al., 2017) and (b) an understanding of how distress changes and unfolds over time (Prenoveau, 2016).

Measurement Invariance

To measure distress during COVID-19, researchers have relied on established mental health scales such as the General Health Questionnaire (GHQ; Goldberg & Williams, 2000). The GHQ-12 assesses general psychological distress in the past 2 weeks via symptoms of common mental health problems such as anxiety, depression, somatic symptoms, and social dysfunction, by using a 4-point Likert scale, leading to scores between 0 and 36 (Gnambs & Staufenbiel, 2018). Given good reliability (internal consistencies between .79 and .91; retest reliability between .68 and .84) and validity of its scores (e.g., sensitivity of .84 and specificity of .79, good convergent validity), the GHQ-12 is widely used in clinical practice and epidemiological research to index psychological distress (see Gnambs & Staufenbiel, 2018, for a meta-analysis). Measurement invariance (MI) of the GHQ-12 has been confirmed in adults across clinical and nonclinical populations (Fernandes & Vasconcelos-Raposo, 2013), different ethnic groups (Bowe, 2017), different cultures (Romppel et al., 2017), gender (Shevlin & Adamson, 2005), and time (Mäkikangas et al., 2006). The GHQ-12 quantifies distress via symptoms like "have you recently been able to enjoy your normal day-to-day activities?" and "have you recently felt you were playing a useful part in things?". However, during the pandemic, responses to these items may reflect lifestyle constraints due to virus mitigation strategies as opposed to the intended psychological distress. Both interpretation shifts and more distressed response rates may result in floor/ceiling effects for items, rendering them unreliable discriminators of distress. It is therefore important to discern whether changes in psychological distress are indeed attributable to changes in the latent construct or an artifact of such measurement properties (Liu et al., 2017). If, over time, symptoms relate differently to the construct (i.e., different factor loadings) or item response categories have a different meaning (i.e., different item thresholds), changes in the resulting psychological distress scores may not reflect true changes in the latent construct, making findings difficult to interpret. In a prior study, sensitivity analyses led to equivalent results when excluding the item "enjoying day-to-day activities" from GHQ analyses during the pandemic (Pierce et al., 2021). Nonetheless, it is important to test MI to identify whether consistent measurement models of GHQ exist over time prior to and during the COVID-19 pandemic.

MI is established with increasingly constrained models that are consecutively tested against the prior model. First, in the test of *configural MI*, the factor structure (number of factors) is constrained to be equivalent across time points. According to a meta-analysis, the GHQ-12 is essentially unidimensional (Gnambs & Staufenbiel, 2018). Multidimensional factor solutions are mainly attributable to some items being negatively phrased and some being positively phrased, which can be addressed by allowing their covariances to vary freely in a unidimensional model. However, this factor solution needs to hold over time during COVID-19 to allow for meaningful comparisons of GHQ-12 total scores. Second, in the test of *weak/metric MI*, factor loadings are additionally constrained to be equal across time to investigate whether the items relate to the latent depression trait in the same way across time points. For instance, if most people indicate they do *not enjoy day-to-day activities* during the pandemic, this item may not be reflective of underlying mental distress and may have lower factor loadings compared to before the pandemic. Third, in the test of *strong/scalar MI*, item thresholds (for ordered-categorical data) are additionally constrained to be equal across time to discern whether the response categories relate to the latent construct in a comparable way. Thresholds for certain items may be lower during the pandemic; for example, more people may readily indicate *not enjoying activities* compared to before the pandemic. Last, in the test of *unique factorial MI*, residual variances of the items are additionally constrained to be equal over time. Unique factorial MI needs to be demonstrated in longitudinal MI with ordered-categorical indicators to ensure that any changes in the means or covariances of the observed scores reflect changes in the underlying latent construct (Liu et al., 2017). Only if these strict assumptions hold can we conclude that changes in composite GHQ scores over time are indicative of genuine changes in psychological distress prior to and during COVID-19. This will enable a more confident interpretation of prior (Ellwardt & Präg, 2021; Pierce et al., 2021), present, and future findings assessing change in GHQ sum scores around the time of the pandemic.

Time-Invariant and Time-Varying Variance

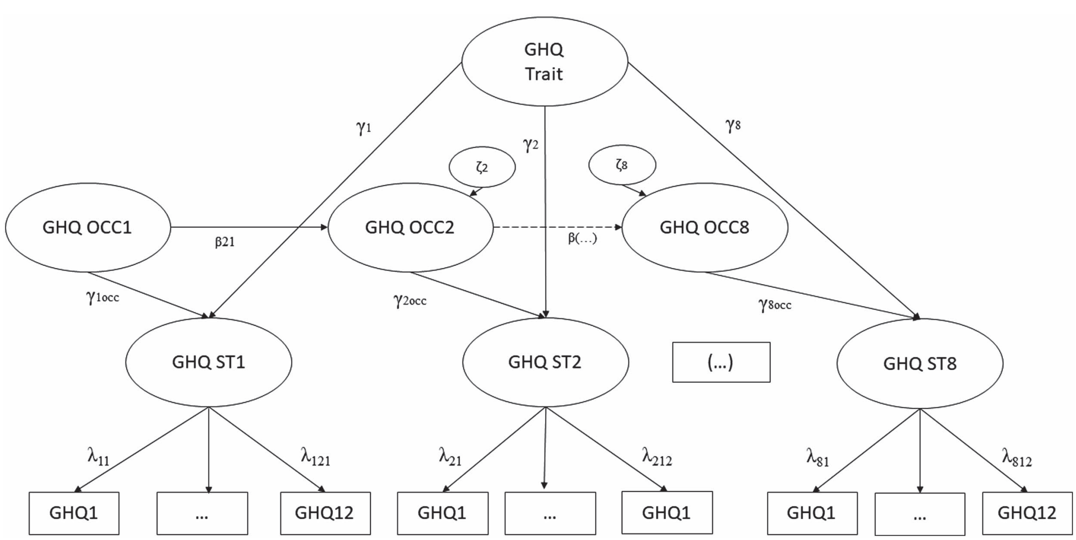
After a measure is shown to be longitudinally invariant, for a more fine-grained understanding of the latent construct over time, one can investigate how much variance is attributable to a chronic *time-invariant* component (trait-like variance: an individual's long-term average in a given sample) and a *time-varying* component (state variance: signifying transient, changing aspects; Prenoveau, 2016). While psychological distress increased during the onset of the COVID-19 pandemic (Pierce et al., 2020), it is possible that a similar rank-order of individuals would be observed, reflecting relative interindividual consistency over time (time-invariant variance). Conversely, the rank-order may change dramatically during COVID-19 if, irrespective of prior symptoms, some individuals develop more psychological distress while others do not (time-varying variance). Such changes in the rank-order may be more pronounced during national lockdowns compared to lower levels of restriction during the pandemic. If the different variance sources remain entangled, investigating risk and protective factors may result in misleading interpretations (Prenoveau, 2016). Distinctively examining the time-varying proportion of a construct enables researchers to investigate which part of the variance is changing over time (Prenoveau, 2016). From a practical perspective, it is important to target this fluctuating part of the variance because it is potentially tractable. Specific to COVID-19, if a large part of variance in distress is transient and related to lockdown restrictions or financial difficulties (for example), it is advantageous to pinpoint the timing of maximal influence of these factors. Such knowledge can elucidate key factors for time-specific intervention, which could be tested and implemented in subsequent lockdowns or future pandemics.

The trait–state–occasion (TSO) model allows the partitioning of these different variance sources (Cole et al., 2017; Prenoveau, 2016). After measurement invariance is established, the underlying latent factors of psychological distress at each wave are divided into one overarching *latent trait factor* across all waves, with a *latent occasion factor* for each wave. The latent trait factor is a higher order latent variable that contains the time-invariant proportion of a construct, the stable interindividual differences across time in a given sample (Figure 1). Thus, the latent trait factor may also be referred to as "time-invariant variance." The latent occasion factors represent the time-varying variance of a given occasion and indicate the relative standing of individuals at each wave. Thus, the latent occasion factors may also be referred to as "time-varying variance." To account for the fact that there are other unspecific carryover effects (e.g., adjacent time points are often more strongly related than others), the model additionally contains *autoregressive pathways* that account for this part of the variance in the data.

Predictors of the Variance Components

Investigating risk and protective factors associated with the different variance components at different phases of the pandemic can provide important targets for intervention. By disentangling the variance sources, one can test whether established predictors of psychological distress during COVID-19 will also be reflected in these models. Using fixed-effect models (which do not differentiate the source of variance), former research identified confirmed COVID-19 cases, local lockdowns, and problems paying bills as time-dependent variables that led to subsequent declines in mental health (Pierce et al., 2021). In latent trajectory analyses, groups with consistently poor mental health were more likely to have preexisting psychiatric diagnoses or ill-health (Ellwardt & Präg, 2021). Social distancing measures during the pandemic led to a decline of in-person contact, which exacerbated loneliness (Bu et al., 2020). Greater loneliness may thus predict changes in psychological

Figure 1
TSO Model of the General Health Questionnaire



Note. ST = state factor; OCC = occasion factor; γ = trait factor loading; γ_{occ} = occasion factor loading; ζ = occasion-specific latent residual variance; λ = factor loadings for reasons of space, this model is only illustrated for three occasions and three indicator variables. β = autoregressive pathway between occasions; TSO = trait–state–occasion; GHQ = General Health Questionnaire.

distress especially during lockdowns. In contrast, social support available from living with a partner may protect mental health at a time where in-person social contacts are scarce (Pierce et al., 2021). Further, life satisfaction predicted psychological resilience during the pandemic (Karataş & Tagay, 2021) and thus may be inversely associated with time-varying distress variance during COVID-19. Financial difficulties, which hit many people unexpectedly during COVID-19, may account for individual differences in changes of psychological distress (i.e., time-varying variance; Chandola et al., 2022). Despite COVID-19 infection risks increasing with age, younger age has been associated with greater increases in psychological distress during COVID-19 (Xiong et al., 2020). A negative association between age and time-varying variance could reflect a greater stress reactivity in younger people, which may be exacerbated by the greater impact of COVID-19-related restrictions on this age group. Investigating the effects of gender appears critical because women displayed more mental health problems during COVID-19, especially at the onset of the pandemic (Pierce et al., 2021). This could reflect higher associations of gender with time-varying variance when governmental restrictions were high. While there was an increase in mental health concerns among ethnic minority groups in the United States (Czeisler, 2020), evidence was more mixed in the United Kingdom (Daly et al., 2022, for subgroup analysis of the present sample).

The Present Study

We examined measurement invariance and interindividual consistency of psychological distress and associated risk factors in the general U.K. population during COVID-19. We conducted secondary analyses using the “Understanding Society, the U.K. Household Longitudinal Study (UKHLS)” data set from 2015 to March 2021 (University of Essex, Institute for Social and Economic Research, 2019). Analyses were conducted on (a) five annual survey waves before COVID-19 (2014–2020) plus the first COVID-19 survey wave and (b) eight monthly or bimonthly survey waves during COVID-19 (April 2020–March 2021). We specifically investigated (a) longitudinal MI of GHQ, and if this was established, we proceeded with TSO modeling to elucidate, (b) time-invariant and time-varying variance components of distress, and (c) predictors of these different variance components. Predictors were based on previously established associations with GHQ during and prior to the COVID-19 pandemic (Pierce et al., 2021). Should GHQ exhibit longitudinal MI, we anticipated that variance would disaggregate into differing levels of time-invariant and time-varying components. Given that there is a lack of knowledge about how predictors relate to time-invariant and time-varying variance components, we refrained from formulating specific hypotheses regarding these analyses.

Method

Participants and Design

The UKHLS sample is representative of the U.K. population. It comprises clustered, stratified samples of households in England, Scotland, and Wales and a nonclustered, systematic random sample in Northern Ireland. Areas with migrant and ethnic minority populations were oversampled. Questionnaires were available in English and Welsh. Annual data collection on UKHLS commenced in 2009; interviews before the pandemic were predominantly conducted in person; the monthly COVID-19 data collection shifted to online or phone interviews during the pandemic. UKHLS is funded by the Economic and Social Research Council and different government departments to understand the experiences of the U.K. population over time with respect to health and socioeconomic factors. Scientific leadership of UKHLS is driven by the Institute for Social and Economic Research, University of Essex. Ethical approval for data collection was given by the University of Essex Ethics Committee (ETH1920-1271). During COVID-19, since April 2020, the survey has been available in an online format, with data collected every 1–2 months in the last week of the month in April (Wave 1), May (Wave 2), June (Wave 3), July (Wave 4), September (Wave 5), November 2020 (Wave 6), January 2021 (Wave 7), and March 2021 (Wave 8) (Table 1). The monthly data collections as part of the COVID-19 study were a subset of the annual main study but both included the GHQ. In the present study, for those who had completed at least 1 COVID-19 wave, $n = 17{,}761$, we analyze the five annual waves of the main survey completed before COVID-19 (2014–2020) and the eight monthly survey waves of the COVID study that were collected during COVID-19 (April 2020–March 2021). Table 1 shows the demographic details of participants and sample size at each wave. Invitations to participate in the COVID-19 study were sent to the 42,330 household members that could be contacted from pre-COVID-19 waves. Of these invited panel members, during COVID-19 waves there were a maximum of 17,761 participants (April 2020, 42.0%) and a minimum of 11,968 (January 2021, 28.3%). Study details have been described elsewhere (Pierce et al., 2021).

Transparency and Openness

All data are openly available to researchers via the U.K. Data Service. This study was not preregistered. Data, program code, and methods used are cited. The sample size was determined by the sampling strategy of the Understanding Society study (University of Essex, Institute for Social and Economic Research, 2019). R analysis code can be found in the OSF: https://osf.io/z6wfb/?view_only=7acaffe491124d2d8fedbd4e4addb05f.

Measures

The *GHQ-12* is a unidimensional 12-item questionnaire assessing mental distress in the past 2 weeks (Goldberg & Williams, 2000) on a 4-point Likert scale, leading to total scale scores from 0 to 36. The GHQ-12 is well-validated across general adult populations (Gnambs & Staufenbiel, 2018). For this Likert-type scoring method, scores above 11 are indicative of caseness of common general psychiatric disorders including depression or generalized anxiety disorder when tested against clinical interviews (Goldberg et al., 1997; Ruiz et al., 2017). Internal consistencies for all waves were excellent (Supplemental Table 1).

Predictor variables of time-invariant and time-varying variance were self-reported. Psychiatric diagnosis was assessed with a dichotomous question at the last wave before COVID-19, “Has a doctor or other health professional diagnosed you with a psychiatric illness?” (no/yes). We also used a dichotomous item asking participants whether they had a “long-standing illness or disability?” (no/yes). Loneliness assessed how often individuals felt lonely from 1 (*hardly ever or never)* to 3 (*often*) and was assessed at all waves during COVID-19 and the two most recent waves before

Table 1
Understanding Society Demographics for the Last Five Waves Pre-COVID-19 and the Eight Waves During COVID-19

Demographic characteristics	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	Wave 7	Wave 8
Pre-COVID	N = 15,234 (2015/2016)	N = 15,666 (2016/2017)	N = 16,022 (2017/2018)	N = 16,261 (2018/2019)	N = 16,604 (2019/2020)	— (April 2020)	—	—
Gender[a]						—	—	—
Male	6,421 (42.1%)	6,593 (42.1%)	6,736 (42.0%)	6,826 (42.0%)	6,957 (41.9%)	—	—	—
Female	8,813 (57.9%)	9,073 (57.9%)	9,286 (58.0%)	9,435 (58.0%)	9,647 (58.1%)	—	—	—
Age	47.80 (15.51)	48.42 (15.80)	48.96 (16.09)	49.44 (16.45)	50.01 (16.70)	—	—	—
COVID-19	N = 17,761 (April 2020)	N = 14,811 (May 2020)	N = 14,123 (June 2020)	N = 13,754 (July 2020)	N = 12,876 (September 2020)	N = 12,035 (November 2020)	N = 11,968 (January 2021)	N = 12,860 (March 2021)
Gender								
Male	7,411 (41.7%)	6,112 (41.3%)	5,803 (41.1%)	5,713 (41.5%)	5,333 (41.4%)	4,992 (41.5%)	4,928 (41.2%)	5,275 (41.6%)
Female	10,334 (58.2%)	8,699 (58.7%)	8,308 (58.8%)	8,032 (58.4%)	7,532 (58.5%)	7,033 (58.4%)	7,030 (58.7%)	7,397 (58.3%)
Age	50.53 (17.06)	52.50 (16.54)	53.03 (16.49)	53.16 (16.54)	54.06 (16.21)	54.57 (16.13)	54.82 (16.12)	54.16 (16.27)
Ethnicity[b]								
White	15,012 (84.5%)	12,869 (86.9%)	12,362 (87.5%)	11,942 (86.8%)	11,333 (88.0%)	10,634 (88.4%)	10,572 (88.3%)	11,099 (87.5%)
Non-White	2,124 (11.9%)	1,550 (10.5%)	1,399 (10.0%)	1,423 (10.3%)	1,242 (9.5%)	1,146 (9.5%)	1,111 (9.3%)	1,267 (9.9%)
Financial								
Comfortably	5,895 (33.2%)	5,141 (34.7%)	—	4,367 (31.8%)	—	3,704 (30.8%)	—	4,047 (31.9%)
Doing alright	7,166 (40.4%)	6,513 (44.0%)	—	6,341 (46.1%)	—	5,562 (46.2%)	—	5,885 (46.4%)
Just right	2,644 (14.9%)	2,302 (15.5%)	—	2,227 (16.2%)	—	1917 (15.9%)	—	1884 (14.8%)
Quite difficult	655 (3.7%)	440 (3.0%)	—	452 (3.3%)	—	417 (3.5%)	—	365 (2.8%)
Very difficult	256 (1.4%)	180 (1.2%)	—	149 (1.1%)	—	157 (1.3%)	—	113 (1.0%)
Partner								
Yes	12,532 (70.5%)	10,309 (69.6%)	9,883 (70.0%)	9,592 (69.7%)	9,083 (70.5%)	8,451 (70.2%)	8,381 (70.0%)	8,922 (70.4%)

[a] Across waves, 0–12 people chose the response option "prefer not to say." The reason for the other missing values is unknown. [b] Ethnicity was dichotomized because the subgroups were too small for our analyses.

COVID-19. Overall, life satisfaction responses ranged from 1 (*completely dissatisfied*) to 7 (*completely satisfied*) and were assessed at Waves 2 (May 2020), 4 (July 2020), 5 (September 2020), 6 (November 2020), 7 (January 2021), and 8 (March 2021) during COVID-19 and at the last two waves before COVID-19. Financial situation was responded to on a scale from 1 (*comfortably*) to 5 (*very difficult*), assessed at COVID-19 Waves 1 (April 2020), 2 (May 2020), 4 (July 2020), 6 (November 2020), and 8 (March 2021) and all waves before COVID-19. Date of birth (to calculate age), sex, and ethnicity were established prepandemic.

Analytical Strategy

We performed all analyses in R (R Core Team, 2021) with the *lavaan* package (Rosseel, 2012). Owing to different time lags between the waves (annual assessments pre-COVID and monthly or bimonthly assessments during COVID-19), models were specified for two different time intervals. First, we conducted a model of the last five annual waves before COVID-19, plus the first COVID-19 wave to pinpoint whether more time-varying variance was present at the start of COVID-19 compared to prior waves. We call this model the *pre-COVID to 1st-lockdown model*. Second, we modeled eight waves during COVID-19 (April 2020–March 2021). We call this the *COVID-19 model*. For all models, fit was considered good when the comparative fit index (CFI) was above .95 (.90 for acceptable fit) and the root-mean-square error of approximation (RMSEA) below .05 (.08 for acceptable fit; Hu & Bentler, 1999). The covariances between errors of the same indicators were allowed to freely vary to account for method effects (Liu et al., 2017).

Missingness

Attrition across waves was a concern (Table 1). Detailed analyses of "Understanding Society" participants who dropped out during the COVID-19 waves compared to waves prior to COVID-19 can be found elsewhere (see Supplemental Material in Pierce et al., 2020, 2021). Notably, those with missing GHQ-12 data were younger, socioeconomically more deprived, and had higher scores on the GHQ-12 at Wave 1 pre-COVID-19 (2014–2015). To have a reasonable amount of missingness, we only analyzed participants who took part in at least one COVID-19 wave (N = 17,761). Of these participants, at least 85.8% (15,234/17,761) of the data were available for each wave in the *pre-COVID to 1st-lockdown model* compared to the COVID-19 Wave 1 (Table 1). Also, at least 67.4% (11,968/17,761) of the data from these participants were available at each wave of the *COVID-19 model* compared to Wave 1 during COVID-19. To account for the ordinal data, we used the weighted least squares mean and variance adjusted (WLSMV) estimator for all models. Simulation studies show that for MI testing, parameter and standard error estimates are relatively unbiased using WLSMV with missingness up to 50% when the sample size is large (≥1,000; Chen et al., 2020). While simulation studies do not directly address TSO modeling, MI is a necessary precondition of the TSO model (Prenoveau, 2016). Furthermore, when using the WLSMV estimator, if the data are missing at random with respect to certain variables, including these variables in the model will yield unbiased estimates for parameters and their standard errors (Asparouhov & Muthén, 2010). Although the missing patterns point to missing at random mechanisms, we applied the extra dependent variable model that allows inclusion of auxiliary variables to account for attrition within our categorical variables using WLSMV (Graham, 2003). Given the model complexity, we could only include one auxiliary variable associated with missingness in our models without encountering convergence problems. Therefore, we included the GHQ mean scores at Wave 1 pre-COVID-19 (2014–2015) as auxiliary variable because they were consistently associated with missingness across waves. This way, we aimed to reduce bias due to attrition of participants with mental health problems. In factor score regression models, we used full-information maximum likelihood to account for missingness.

Measurement Invariance

For the TSO model, we followed the steps outlined by Prenoveau (2016). First, we needed to establish longitudinal MI to ensure that the same latent construct was measured over time (Liu et al., 2017). We first tested whether the unidimensional GHQ factor structure remained the same over time (configural MI). Despite some studies allowing the covariances of errors among negatively phrased items (Gnambs & Staufenbiel, 2018), we first tested a simple one-factor model without covariances of errors among single items. If this model produced an adequate fit, then further modifications were not necessary. Next, factor loadings and item thresholds of the same indicators were constrained to be the same over time (scalar MI). For categorical MI, the steps of constraining factor loadings and thresholds to be equal are usually combined (Chen et al., 2020). Last, we set the item residuals to equity over time (unique factor MI). To demonstrate MI, the ΔCFI should not exceed .010 and the ΔRMSEA .007 (Neufeld et al., 2022).

TSO Modeling

Next, we specified the TSO model with the MI constraints for the highest established level of MI (Prenoveau, 2016, Figure 1): The 12-GHQ items with the MI constraints per wave constituted the state factors that correspond to the number of waves (6 waves for Model 1 and 8 waves for Model 2). A higher order latent trait factor was then introduced with factor loadings to each of the latent state factors, representing the time-invariant proportion of the variance. For each wave, a latent occasion factor was specified on which the state factor loaded, representing the time-varying proportion of the variance. Here, to put the metric on the occasion factor, the coefficients of the state factors on the occasion factors were set to one. To divide the variance into time-variant and noninvariant proportions, the residual variances of the state factors were set to zero. Then, to account for the fact that the occasion factors tend to perpetuate themselves, we specified autoregressive pathways between the occasions. This way, the variance from the second occasion onward can be disentangled into three sources: time-invariant proportion, time-varying proportion, and variance explained by the autoregressive pathway. The squared standardized loadings γ^2 (factor loadings of the states on the trait factor) provide the proportion of variance that is attributable to the higher order GHQ trait (time-invariant proportion). Calculating $1 - \gamma^2$ provides the variance that is occasion-specific (time-varying). From the second wave onward, the occasion-specific variance can be further disentangled by using the squared standardized autoregressive pathways β^2 between two consecutive waves. The variance in the occasion factor that is explained by the autoregressive pathway can be calculated by multiplying the occasion-specific variance with the squared standardized autoregressive pathway. The leftover (1 − time-invariant proportion—autoregressive proportion) constitutes the time-varying component. To ensure model identification, we constrained the autoregressive pathways and the residual variances from the second to last wave to equity, respectively (Prenoveau, 2016).

The TSO model was tested against simpler models to ensure that this more complex model provided the best fit to the data: (a) a trait stability model, where autoregressive paths were removed, indicating that there are no unspecific carryover effects and that the variance is fully captured by the trait and occasion factors; (b) an autoregressive-only model, where the latent trait factor was removed, indicating that there is no stability in the rank-order over time that can be captured by a time-invariant factor; and (c) equal trait-to-state loadings model, where factor loadings of the GHQ trait were set to equity to test whether the occasion variance was equal across measurement occasions. Each of these models was tested against the full TSO model, with a nonsignificant scaled chi-square difference test indicating model equivalence, and that the simpler model should be used (Prenoveau, 2016).

Prediction Models

When the best fitting model was identified, we used different variables to predict the different proportions of the variance (i.e., time-invariant variance, and for each wave occasion-specific time-varying variance). Due to the model complexity, we first derived factor scores for the different variance components with the empirical Bayes method (Muthén, 1998). We then tested loneliness, living with a partner, life satisfaction, financial difficulties, ill-health (only assessed during the pandemic), age, gender, and ethnicity as simultaneous predictors of the standardized factor scores of the different variance components. For the time-varying variance component, multiple regression models included all predictors at the previous wave to predict the occasion variance of the next wave (with the exception of the first wave, when we used contemporaneous associations). When the variable was not assessed at the previous wave, we used the wave before this wave. For the time-invariant variance, we tested one multiple regression model with all predictors of time-invariant variance from the first COVID-19 wave because so many predictors over time could be used to predict time-invariant variance (e.g., life satisfaction at pre-COVID Waves 1–8 and loneliness at pre-COVID Waves 1–8). However, this one wave had the most unique variables available of all the waves. In total, we ran 16 regression models, as we predicted time-varying variance for six waves for the pre-COVID to 1st-lockdown model and for eight waves in the COVID-19 model in addition to predicting the time-invariant variance for both models. A positive association between a predictor variable with trait distress variance suggests higher GHQ-12 scores relative to the trait sample mean. A positive association between a predictor variable with occasion-specific distress variance indicates higher GHQ-12 scores relative to the occasion-specific sample mean at a given occasion.

Results

Descriptive Statistics

Descriptive statistics of the GHQ composite scores, loneliness item, and GHQ internal consistencies are provided in Supplemental Table 1. Across waves, internal consistencies were excellent (0.90+). Figure 2 shows mean GHQ changes across time, including where consecutive means significantly differ.

Measurement Invariance

All MI models displayed good model fit according to the CFI and acceptable-to-good fit according to the RMSEA for the one-factor model, with no error covariances (Table 2). We could establish the highest level of invariance with unique longitudinal MI across both time intervals as indicated by no deterioration in model fit when models were increasingly constrained. All factor loadings were good at all measured waves (>.40), indicating all items loaded well on the distress construct at all waves. During national lockdowns or partial lockdowns (April, May, June, and November 2020, January and March 2021), the items, *have you recently been able to enjoy your normal day-to-day activities?* and *have you recently felt you were playing a useful part in things?* had equivalent loadings to other waves (range of .48−.55 over all waves), suggesting that these items did reflect psychological distress as opposed to lifestyle constraints due to virus mitigation strategies.

TSO Modeling

Each model had good fit according to the CFI and acceptable to good fit according to the RMSEA. For both models, removing the autoregressive pathway in the trait-stability model and removing the GHQ-trait factor in the autoregressive-only model led to a significant decrement in model fit (Table 2). Setting all loadings equal so that the variance components were equal also led to a decrement in model fit. We therefore continued with the TSO model.

Variance Proportions

Pre-COVID to 1st-Lockdown Model

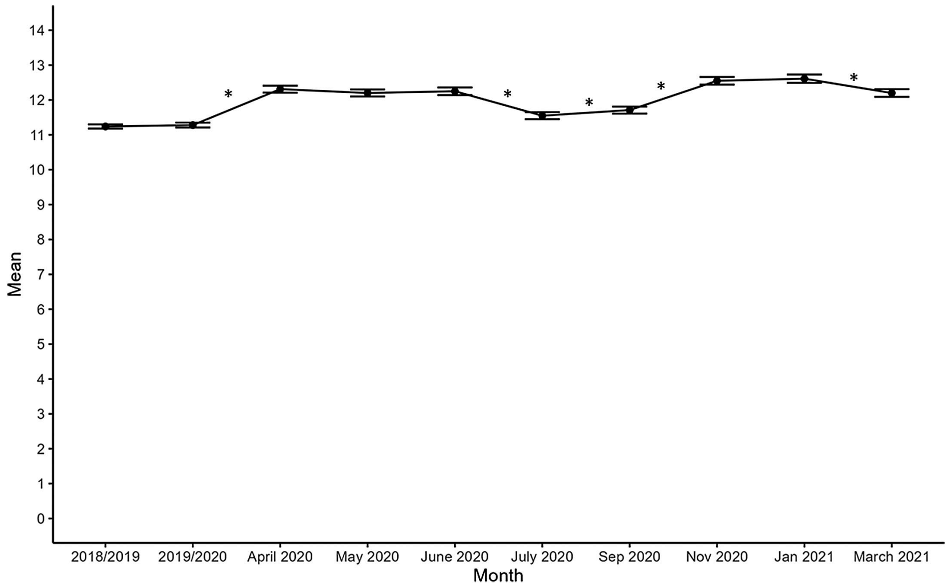
As depicted in Figure 3 and Table 3, for the model of the five annual waves before COVID-19 plus the first COVID-19 assessment (April 2020), the time-invariant distress variance ranged from 37.6% (April 2020) to 52.2% (2018/2019). The time-varying variance ranged from 46.8% (2018/2019) to 61.2% (April 2020). Time-varying variance was statistically higher in April 2020, as indicated by nonoverlapping confidence intervals (Table 3). Inspection of confidence intervals indicates the increase in population estimates of time-varying variance from pre-COVID waves to April 2020 ranged from 4.9% to 17.2% based on examination of upper and lower bounds of the confidence intervals. Approximately 1% of the variance was explained by the autoregressive paths.

Sensitivity Analyses

After we ran this model, we conducted sensitivity analyses to gauge whether models that only included equal time lags influenced the variance estimates of the April 2020 assessment. We aimed to ensure that the high levels of time-varying variance in April 2020 were not attributable to the different timeframes used in the models. To this end, we extended the pre-COVID to 1st-lockdown model and added the last COVID-19 assessment in March 2021. In addition, we tested a model only including Wave 3 (2017/2018) and Wave 4 (2018/2019) before COVID-19 and April 2020 and

Figure 2

Mean Changes in General Health Questionnaire (Mental Distress) in the Understanding Society Data Set Before and During the COVID-19 Pandemic



Note. Key COVID-19 pandemic dates: April 2020–lockdown 1; July 2020–restrictions eased; November 2020–lockdown 2; January 2021–lockdown 3. All significant changes are a small effect size except July–September 2020 and January–March 2021, which are negligible (see Supplemental Material 1, for effect sizes). Error bars represent 95% CI. Scores above 11 indicate clinical caseness (Goldberg et al., 1997; Ruiz et al., 2017). CI = confidence interval.

* Signifies changes at $p < .001$ determined using t tests.

March 2021. These were the most plausible models because the time gap between the first and the last COVID-19 assessment approximately mirrored the time gap of 1 year for the first five assessments before COVID-19. We provide further rationale and model results in the Supplemental Material. Results were in line with the pre-COVID to 1st-lockdown model with the highest level of time-varying variance during April 2020 (60.5% and 63.8%, respectively, per model), compared to the assessments before COVID-19 (highest

Table 2

Measurement Invariance and Latent State–Trait Models of the GHQ Scores Over Time

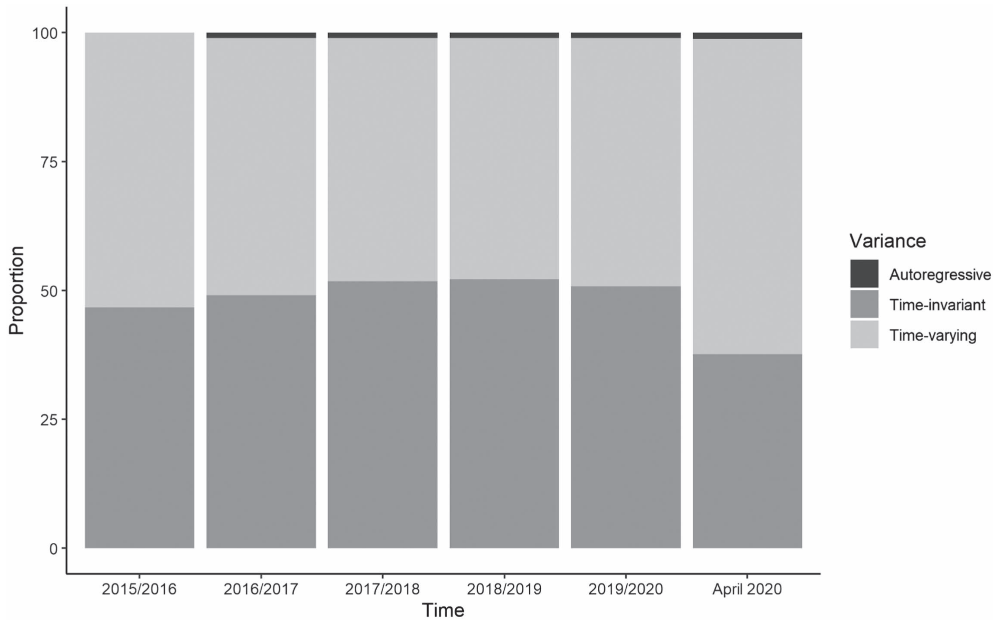
Invariance	$\chi^2(df)$	CFI	RMSEA	ΔCFI	ΔRMSEA
Pre-COVID to 1st-lockdown					
Configural	135,436 (2,351)	.975	.061		
Scalar	138,539 (2,526)	.975	.061	.000	.000
Strict	164,414 (2,590)	.970	.058	.005	.003
COVID-19 model					
Configural	254,186 (4,217)	.978	.074		
Scalar	255,885 (4,440)	.978	.076	.000	.001
Strict	272,510 (4,525)	.976	.075	.002	.001

TSO models	$\chi^2(df)$	CFI	RMSEA	$\Delta\chi^2$	Δdf	p
Pre-COVID to 1st-lockdown						
1. TSO full	163,685 (2,602)	.970	.054			
2. Trait stability model	168,749 (2,603)	.969	.055	689	1	<.001
3. Autoregressive-only	351,597 (2,612)	.935	.070	8,993	6	<.001
4. Equal trait loadings	178,476 (2,606)	.967	.057	478	4	<.001
COVID-19 model						
1. TSO full	274,476 (4,548)	.976	.063			
2. Trait stability model	292,930 (4,549)	.974	.064	846	1	<.001
3. Autoregressive-only	347,417 (4,557)	.970	.063	2,316	9	<.001
4. Equal trait loadings	276,944 (4,554)	.976	.065	59	6	<.001

Note. GHQ = General Health Questionnaire, measuring mental distress; CFI = comparative fit index; RMSEA = root-mean-square error of approximation; df = degrees of freedom; TSO = trait–state–occasion.

Figure 3

Variance Proportions of General Health Questionnaire (Mental Distress) in the Pre-COVID to 1st-Lockdown Waves of the Understanding Society Data Set



level of time-varying variance was 54.5%). Notably, time-invariant variance during March 2021 was at the same levels as the assessments before COVID-19 (51% and 51.3%).

COVID-19 Model. For the eight waves during COVID-19, the time-invariant variance ranged from 59.3% (April 2020) to 67.7% (July 2020). The time-varying variance ranged from 29.0% (July 2020) to 40.7% (April 2020). Again, time-varying variance was statistically higher in April 2020 with confidence intervals indicating the difference in population estimates of time-varying variance in April 2020 compared to the other COVID-19 waves ranged from 4.8% to 14.0% based on examination of upper and lower bounds of the confidence intervals. The autoregressive pathways explained between 3.3% (July 2020) and 3.6% (September 2020) of the variance (Table 3 and Figure 4).

Prediction Models

Pre-COVID to 1st-Lockdown Model

Time-Invariant Distress Variance. Given so many predictors could be used to predict time-invariant variance (e.g., life satisfaction at pre-COVID Waves 1–8 and loneliness at pre-COVID Waves 1–8), we tested one multiple regression model with all predictors of time-invariant variance from the first COVID-19 wave (Table 4). In this model, all variables apart from ethnicity were significant.

Time-Varying Distress Variance. Table 4 depicts the multiple regression models of the time-varying variance components with the different predictors. Each time-varying variance at a given wave was predicted by variables from the former wave (but contemporaneous if the variable was not assessed previously). Briefly, gender predicted time-varying variance at the 2016–2017 wave and the first COVID-19 wave (April 2020). Financial difficulties predicted the next wave of time-varying variance at three waves from 2015 to 2018, and also 2019/2020, but not at the first COVID-19 wave. However, as financial status may have changed related to the pandemic, we also report this contemporaneous association, which was significant. For all waves where loneliness and life satisfaction were available, these predicted time-varying variance at the next wave. Younger age at the prior wave predicted time-varying variance during April 2020.

COVID-19 Model

Time-Invariant Distress Variance. Again, for the time-invariant variance, we tested one multiple regression model of time-invariant variance with all predictors from the first COVID-19 wave (April 2020, but life satisfaction from May 2020; Table 5). All variables were significant.

Time-Varying Distress Variance. Table 5 presents multiple regression models. As for the pre-COVID to 1st-lockdown model, each time-varying variance at a given wave is predicted by variables from the former wave; only Wave 1 depicts contemporaneous associations. In May 2020 and June 2020, psychiatric diagnosis was associated with time-varying variance. Loneliness predicted time-varying variance at all subsequent waves. Female gender predicted time-varying variance at April 2020, May 2020, November 2020, and January 2021 waves. Financial difficulties predicted subsequent time-varying variance at all waves except for November 2020. Lower life satisfaction was consistently associated with subsequent time-varying variance following the waves when this was assessed. Age was negatively associated with subsequent time-varying variance at the June 2020 wave.

Discussion

Drawing on national, probability-sampled longitudinal data in adults from eight waves during the COVID-19 pandemic (April

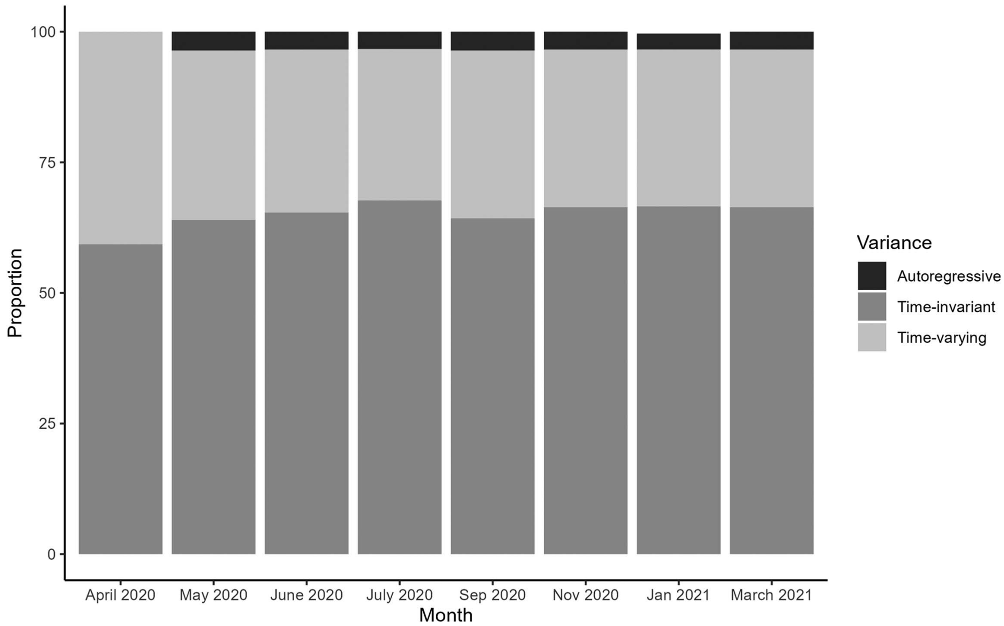
Table 3
Proportion of Variance Across Waves of the General Health Questionnaire (Mental Distress) TSO Models

Variance component	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	Wave 7	Wave 8
Pre-COVID to 1st-lockdown	2015/2016	2016/2017	2017/2018	2018/2019	2019/2020	April 2020		
Trait variance	46.7%	49.1%	51.8%	52.2%	50.8%	37.6%	—	—
95% CI	[45.2; 48.0]	[47.7; 50.5]	[50.4; 53.1]	[50.8; 53.4]	[49.4; 52.1]	[35.8; 39.1]		
Time-varying variance	53.3%	49.9%	47.2%	46.8%	48.2%	61.2%	—	—
95% CI	[54.8; 52.0]	[51.2; 49.5]	[48.6; 46.0]	[48.2; 45.7]	[49.6; 46.9]	[62.9; 59.7]		
Autoregressive	—	1.0%	1.0%	1.0%	1.0%	1.2%	—	—
95% CI	—	[1.1; 1.0]	[1.0; 0.9]	[1.0; 0.9]	[1.0; 1.0]	[1.3; 1.2]		
COVID-19 model	April 2020 (lockdown 1[a])	May 2020	June 2020	July 2020	September 2020	November 2020 (lockdown 2[a])	January 2021 (lockdown 3[a])	March 2021
Trait variance	59.3%	64.0%	66.4%	67.7%	64.3%	66.4%	66.6%	66.4%
95% CI	[58.2; 61.9]	[62.7; 65.3]	[65.1; 67.7]	[66.4; 69.1]	[62.9; 65.8]	[64.9; 67.7]	[65.3; 68.1]	[65.1; 67.9]
Time-varying variance	40.7%	32.4%	30.2%	29.0%	32.1%	30.2%	30.1%	30.2%
95% CI	[41.8; 38.1]	[33.6; 31.2]	[31.4; 29.0]	[30.2; 27.8]	[33.4; 30.8]	[31.6; 29.0]	[31.2; 28.7]	[31.4; 28.9]
Autoregressive	—	3.6%	3.4%	3.3%	3.6%	3.4%	3.3%	3.4%
95% CI		[3.7; 3.5]	[3.5; 3.3]	[3.4; 3.1]	[3.7; 3.4]	[3.5; 3.3]	[3.5; 3.2]	[3.5; 3.2]

Note. Pre-COVID: Waves 1–5 before COVID-19; pre–first model (pre-COVID to 1st-lockdown): Waves 1–5 before COVID-19 + Wave 1 during COVID (Wave 6 in this model, collected April 2020), COVID-19 = Wave 1–8 during COVID-19. TSO = trait–state–occasion; CI = confidence interval.
[a] In the first and third lockdowns, children were homeschooled (except for key worker children). People were only allowed outside for traveling to work (that could not be done at home), exercise once a day, and for essential shopping or medicines. In the second lockdown, children were attending school but extracurricular activities were stopped. People were instructed to stay at home where possible.

Figure 4

Variance Proportions of General Health Questionnaire (Mental Distress) in the COVID-19 Model From Understanding Society Data



2020–March 2021) and five prepandemic waves (2015–2020), we investigated (a) longitudinal MI of mental distress, (b) time-invariant and time-varying variance components, and (c) predictors for these different variance components during the pandemic.

Measurement Invariance

These findings are the first to establish unique longitudinal factorial MI of the GHQ throughout and prior to the pandemic. Our findings support the GHQ one-factor structure, which is consistent with prepandemic adult populations (meta-analysis: Gnambs & Staufenbiel, 2018). This factor structure demonstrated the highest level of unique factorial MI across all waves, both during and prior to the pandemic. This provides evidence that change in distress over time measured with the GHQ-12 can be interpreted as "true" change in latent mental distress (Liu et al., 2017), despite the many changes occurring throughout the pandemic. Potential interpretation shifts in certain items during the pandemic (e.g., "enjoying day-to-day activities") were not supported, as all item factor loadings remained similar across time. Furthermore, as MI also held when data collection moved from in-person to online, these data collection changes also did not appear to have altered responses. Establishing longitudinal unique MI in GHQ increases confidence that prior findings of increased distress in adults during the

Table 4

Pre-COVID to 1st-Lockdown Data: Multivariable Associations of the Time-Varying (TV) and Time-Invariant (TI) GHQ Variance Components With All Predictors

Predictors	TI	TV1 (2015/2016)	TV2 (2016/2017)	TV3 (2017/2018)	TV4 (2018/2019)	TV5 (2019/2020)	TV6 (April 2020)
Adjusted R^2	.28	.02	.01	.01	.10	.03	.05
Loneliness	.47***	—	—	—	.24***	.09***	.08***
Gender (female)	.18***	.04	.07**	.03	−.03	.03	.14***
Financial hardship	.23***	.17***	.12***	.11***	.01	.06***	.02 (.07***)[a]
Age	−.01***	.01†	.01†	.00	−.01†	.00	−.01***
Ethnic minority	.06	.00	.00	.00	.00	−.01	.00
Life satisfaction	−.15***	—	—	—	−.17***	−.07***	−.04***
Psychiatric diagnosis	.51***	—	—	—	—	—	.06
Ill-health	.20***	—	—	—	—	—	.01

Note. GHQ = General Health Questionnaire. Apart from Wave 1 (cross-sectional financial difficulties) and Wave 4 (cross-sectional loneliness and life satisfaction), the previous (possible) wave is used to predict time-varying variance. Time-invariant variance is predicted by first COVID-19 wave variables (Wave 6 here). Estimates are fully standardized.

[a] Association with concurrent financial hardship.

** $p < .01$. *** $p < .001$. † $p < .05$.

Table 5
COVID-19 Model: Multivariable Associations of the Time-Varying Variance (TV) and Time-Invariant (TI) GHQ Variance Components With All Predictors

Predictors	TI	TV1 (April 2020)	TV2 (May 2020)	TV3 (June 2020)	TV4 (July 2020)	TV5 (September 2020)	TV6 (November 2020)	TV7 (January 2021)	TV8 (March 2021)
Adjusted R^2	.36	.15	.08	.06	.03	.02	.03	.03	.03
Loneliness	.60***	.61***	.39***	.25***	.16***	.17***	.22***	.22***	.16***
Gender (female)	.20***	.14***	.07**	.01	−.04	.03	.06†	.06†	−.04
Age	−.04**	.00	.00	−.01***	−.00	−.00	.00	.00	.00
Ethnic minority	.12**	−.02	−.02	.09	−.01	.07	.07	.08	.04
Financial hardship	.21***	.11***	.10***	.04†	.07***, a	.04†	.03[a]	.04†	.05**, a
Partner (no/yes)[b]	.12***	.21***	.04	.01	.01	.04	.07†	.04	.01
Life satisfaction	−.14***	—	—	−.07***	−.04***, a	−.03***	−.05***	−.04***	−.05***
Psychiatric diagnosis	.37***	.08	.14**	.13†	.09	.10	.05	.05	.01
Ill-health	.12***	.03	.01	.01	.02	.03	.02	.04	.04

Note. GHQ = General Health Questionnaire. Apart from Wave 1 (cross-sectional), the previous (possible) wave is used to predict time-varying variance. Time invariant variance is predicted by the first COVID-19 wave (and life satisfaction for Wave 2). Estimates are fully standardized.
[a] Not previous wave but wave before this because this variable was not assessed at previous wave. [b] The association with having a partner at Waves 1 and 6 and for the time-invariant variance was considered spurious because of an implausible sign flip compared to univariate associations, potentially imposed by the correlation of having a partner and loneliness, $r = .31$. Hence, we do not interpret this association.
† $p < .05$. ** $p < .01$. *** $p < .001$.

pandemic (Ellwardt & Präg, 2021) are not biased by measurement changes over time. In fact, for ordered-categorical indicators, unique factorial MI needs to be demonstrated to ensure that any changes in the means or covariances of the observed scores can be meaningfully interpreted (Liu et al., 2017). Accordingly, the use of composite scores over time as used in previous research seems justified to indicate changes in psychological distress during COVID-19 (Pierce et al., 2020). These findings support the robustness of GHQ in assessing mental distress in future general population adult samples, which could be particularly informative for policymakers should similar stressors be faced in the future. This is important given that other measures such as the Children's Depression Inventory could not demonstrate scalar measurement invariance for symptom assessment before and after the outbreak of the pandemic (Olino et al., 2022). Finally, findings point to the necessity of investigating longitudinal MI of other measures that are commonly applied to quantify psychological distress during the pandemic (e.g., Shevlin et al., 2022).

TSO Modeling

To more fully understand the longitudinal change in distress, we disentangled the variance components of the underlying construct over time (Prenoveau, 2016). Time-varying variance increased during the first measurement wave of COVID-19 (April 2020) compared to the five waves before COVID-19 (2015–2020). This demonstrates that increases in mental health symptoms at the onset of the pandemic (Pierce et al., 2020) were associated with rank-order changes in the construct. In other words, irrespective of prior symptoms, some individuals developed more psychological distress while others did not. Our sensitivity analyses, which included only equally spaced waves both pre- and during COVID-19, also revealed that time-varying variance was higher during April 2020 compared to all other waves, bolstering support for our interpretation.

Lower levels of trait-like variance found at the first COVID-19 wave (April 2020) compared to later COVID-19 waves also point to more rank-order changes in distress at the beginning of the pandemic. The high levels of time-varying variance during April 2020 in both models and sensitivity analyses indicate that the first lockdown may have presented a major disruption in the lives of many people leading to distress in different individuals.

While the lowest time-varying variance in distress was found in July 2020, a time with fewer restrictions in the United Kingdom, including "Super Saturday," when restaurants, hairdressers, and pubs were allowed to reopen (Aspinall, 2020), confidence intervals were still overlapping with other assessment points after April 2020. Importantly, time-varying variance in distress did not significantly heighten during subsequent lockdowns, such as in January 2021, which had equivalent restrictions to those imposed at the start of the pandemic. Despite a significant mean increase in distress during lockdowns in November 2020 and January 2021 (Figure 2), the rank-order remained more stable compared to the first lockdown. This suggests that those who were already vulnerable to distress at this point in the pandemic (on a trait-like level) were more likely to display higher levels of mental distress.

One could have expected that the time-varying variance component would be higher in the model where all waves occurred during the pandemic compared to the model that included the five annual waves before COVID-19. However, time periods differed across the two models: The eight COVID-19 waves were obtained within 1 year with a maximum of 2 months between assessments, whereas GHQ was assessed annually in the pre-COVID to 1st-lockdown model. TSO modeling assumes that shorter timeframes reveal less time-varying variance because a shorter period can be better captured by an overarching time-invariant "trait" compared to a time interval of years. Therefore, the time-varying variance during the eight COVID-19 waves needs to be interpreted considering the different time periods in these models (Prenoveau, 2016). In the absence of major events like COVID-19, the time-varying variance in distress may have been even smaller within such a circumscribed period. This is also reflected in the differences in the time-varying variance of the first COVID-19 wave that changed from 61.2% in the pre-COVID to 1st-lockdown model to 40.7% in the COVID-19 model.

Prediction Models

Changes in the relative standing of individuals on the latent distress trait allowed further understanding of the specific predictors of the variance components. In our regression models, we found rather small effects despite statistical significance at $p < .001$. The large sample size allowed for small standard errors, resulting in some small effects becoming significant. Thus, not all predictors may be meaningful in the present study, but as no other study has yet reported on predictors of these partitioned variances, a practically meaningful effect size in this context is unknown. However, more variance in time-invariant variance was explained in multiple regression models compared to time-varying variance. We therefore interpret results based on the consistency of effects over time and across models (both pre-COVID to 1st-lockdown model and the COVID-19 model). In multiple regression models, loneliness, life satisfaction, and financial hardship were the most consistent predictors of time-varying variance.

Loneliness was consistently associated with the trait and time-varying variance components of distress prior to and during the pandemic, consistent with loneliness' reported relationship with mental health (Beutel et al., 2017). Associations with time-varying variance for the first two waves during the pandemic (April and May 2020) were higher than for subsequent waves. This indicates that feeling lonely at the beginning of the pandemic was an important factor that contributed to new onset of psychological distress in people who were presumably affected by social distancing measures (Magson et al., 2021). In multiple regressions, living with a partner was not significant, pointing to the importance of the subjective feeling of loneliness over this more objective indicator of social support. Individuals may typically seek social support from friendship networks and family members not living with them, and thus simply living with a partner may not be sufficient to counteract feelings of distress throughout the pandemic.

Life satisfaction was associated with lower variance for both components of distress in accord with this construct's reported relationship with mental health (Lombardo et al., 2018). Interestingly, associations with time-varying variance had a similar magnitude both prior to and during the pandemic. Therefore, life satisfaction appears to be a general protective factor mitigating distress, as also reflected in life satisfaction's association with trait-like distress. During COVID-19, individuals may have lower life satisfaction for a variety of reasons (e.g., loss of employment, inability to engage in hobbies, being socially less active), and it is important to address specific factors that interfere with life satisfaction.

Financial difficulties predicted time-varying variance consistently throughout the COVID-19 waves studied and during most waves prior to COVID-19. Specific to COVID-19, many individuals encountered financial difficulties for the first time (Chandola et al., 2022), which may have been exacerbated by uncertainties during COVID-19. This is in line with findings that having problems paying bills emerged as time-dependent variable predicting subsequent declines in mental health during COVID-19 (Pierce et al., 2021).

The presence of a preexisting psychiatric diagnosis was associated with trait variance and with time-varying variance in May 2020 and June 2020, as the first COVID-19 lockdown was beginning to ease. In contrast to former studies that did not separate time-invariant and time-varying variance (Ellwardt & Präg, 2021), the effect of preexisting conditions was thus not as central compared to the effects of other predictors. In both models (pre-COVID to Wave 1 COVID and eight COVID-19 waves), women displayed greater time-invariant distress prior to and during the pandemic, greater time-varying variance predominantly at the first COVID-19 wave, and slightly higher time-varying variance in distress during November 2020 and January 2021, in line with findings that women report more anxiety and depression than men (Salk et al., 2017) and with previous studies on this sample that women reported more distress at the onset of pandemic than men (Pierce et al., 2021). Age was associated with time-invariant distress variance and time-varying variance during some waves. However, effects of age and gender were generally very small and thus appear not to be important predictors in the present context. Ill-health and ethnicity were not associated with time-varying variance and seem thus not important in the context of these models.

Although the above associations have been previously reported, they have not been contextualized with respect to their time-invariant and their time-varying variance. Associations were stronger for the time-invariant variance and only low associations have been found for the time-varying variance. Some of the associations with time-varying variance were likely significant because of the large sample, despite reflecting small amounts of explained variance. However, these associations explain variance after stable interindividual differences have been accounted for. Although small, these associations with time-varying variance account for actual changes over time. Predictors of these changes may thus be more amenable to intervention than those which are confounded by trait variance. Likewise, mean-level changes reported here only reached small effect sizes. This is a population-based study, and thus even small effect sizes may be meaningful if they affect many people (Greenberg & Abenavoli, 2017). Despite these mean-level increases, prior to COVID-19, mean GHQ levels were already above clinical thresholds of 11 (Goldberg et al., 1997). Therefore, more research is warranted to contextualize these parameter estimates with regard to clinically significant change (see Daly et al., 2022, for an analysis of this sample).

Strengths and Limitations

Our study has considerable strengths including the sample composition, longitudinal design, and analytical approach. Although studies using this data set have contributed important knowledge by identifying latent trajectories during the pandemic (Ellwardt & Präg, 2021; Pierce et al., 2021), the present study demonstrates that the GHQ measures the same construct throughout the pandemic, decomposes the variance into a stable and fluctuating part, and identifies predictors for these variance components.

There are methodological limitations. First, there is no clear consensus regarding the evaluation of MI for categorical data (Liu et al., 2017). Cutoff criteria for fit indices have not been systematically tested, and the χ^2-statistic may inflate small differences when the sample size is large. The same applies to TSO modeling for which discrepancies in fit indices have not been tested, and thus model comparisons rely on the χ^2-statistic and considerations of overall model fit. It is worth mentioning that the alternative trait stability model and the equal trait loading model had also good model fit in both the pre-COVID to 1st-lockdown and in the COVID-19 model. Indeed, in the pre-COVID to 1st-lockdown model only around one percent of the variance was explained by the autoregressive pathway, while around 3% of the variance was explained by autoregressive pathways in the COVID-19 model. These differences are likely attributable to the shorter periods between assessments in the COVID-19 model that allow for more unspecific carryover effects. While the magnitude of occasion variance was similar across waves in both models, it differed significantly for the first COVID-19 wave in both models, thus providing important information. Despite the more parsimonious alternative models, the full TSO model contained more detailed and crucial information relevant for the present research question.

Attrition was a concern in the present study. Patterns of nonresponse have been described in detail earlier (Pierce et al., 2021). Younger age, lower socioeconomic status, and higher scores on the GHQ-12 at Wave 1 pre-COVID-19 (2014–2015) were associated with attrition. Age, financial difficulties, and psychiatric diagnosis were included in regression models that may thus have been biased by these patterns of missingness. Due to model complexity, we could not adjust for all these variables in the present analyses, and some mechanisms of missingness remain unknown. Therefore, we cannot ensure that our missing at-random assumption holds. Although the GHQ-12 has shown MI across clinical and nonclinical populations (Fernandes & Vasconcelos-Raposo, 2013), it remains unknown whether this is also the case during the pandemic and how this may have affected our TSO modeling. Nonetheless, the sample remained large and nonresponse declined with time during the pandemic assessments (i.e., largest dropout was found for COVID-19 Wave 1 compared to Wave 2). Thus, the level of bias in the data is unlikely to be consequential. Last, general health, psychiatric diagnosis, loneliness, financial hardship, and life satisfaction were all measured with single items, potentially limiting their reliability and construct validity (e.g., assessing financial situation rather subjectively).

Conclusion

The present study indicates that the GHQ-12 reliably assessed changes over time in psychological distress during the COVID-19 pandemic and demonstrates that the beginning of the COVID-19 pandemic was associated with more time-varying variance in distress in adults. The latter suggests that there is a crucial window of time for intervention efforts. Finally, our findings suggest that mitigation efforts should focus on those experiencing loneliness, decreased life satisfaction, and financial hardship, as these factors consistently predicted time-varying distress during the pandemic.

References

Asparouhov, T., & Muthén, B. (2010). Weighted least squares estimation with missing data. *Mplus Technical Appendix*, *2010*(1–10), Article 5.

Aspinall, E. (2020, April 8). *COVID-19 timeline*. British Foreign Policy Group. https://bfpg.co.uk/2020/04/covid-19-timeline/

Banks, J., & Xu, X. (2020). The mental health effects of the first two months of lockdown during the COVID-19 pandemic in the UK. *Fiscal Studies*, *41*(3), 685–708. https://doi.org/10.1111/1475-5890.12239

Beutel, M. E., Klein, E. M., Brähler, E., Reiner, I., Jünger, C., Michal, M., Wiltink, J., Wild, P. S., Münzel, T., Lackner, K. J., & Tibubos, A. N. (2017). Loneliness in the general population: Prevalence, determinants and relations to mental health. *BMC Psychiatry*, *17*(1), Article 97. https://doi.org/10.1186/s12888-017-1262-x

Bowe, A. (2017). The cultural fairness of the 12-item General Health Questionnaire among diverse adolescents. *Psychological Assessment*, *29*(1), 87–97. https://doi.org/10.1037/pas0000323

Bu, F., Steptoe, A., & Fancourt, D. (2020). Who is lonely in lockdown? Cross-cohort analyses of predictors of loneliness before and during the COVID-19 pandemic. *Public Health*, *186*, 31–34. https://doi.org/10.1016/j.puhe.2020.06.036

Chandola, T., Kumari, M., Booker, C. L., & Benzeval, M. (2022). The mental health impact of COVID-19 and lockdown-related stressors among adults in the UK. *Psychological Medicine*, *52*(14), 2997–3006. https://doi.org/10.1017/S0033291720005048

Chen, P.-Y., Wu, W., Garnier-Villarreal, M., Kite, B. A., & Jia, F. (2020). Testing measurement invariance with ordinal missing data: A comparison of estimators and missing data techniques. *Multivariate Behavioral Research*, *55*(1), 87–101. https://doi.org/10.1080/00273171.2019.1608799

Cole, D. A., Martin, J. M., Jacquez, F. M., Tram, J. M., Zelkowitz, R., Nick, E. A., & Rights, J. D. (2017). Time-varying and time-invariant dimensions of depression in children and adolescents: Implications for cross-informant agreement. *Journal of Abnormal Psychology*, *126*(5), 635–651. https://doi.org/10.1037/abn0000267

Czeisler, M. É. (2020). Mental health, substance use, and suicidal ideation during the COVID-19 pandemic—United States, June 24–30, 2020. *MMWR: Morbidity and Mortality Weekly Report*, *69*(32), 1049–1057. https://doi.org/10.15585/mmwr.mm6932a1

Daly, M., Sutin, A. R., & Robinson, E. (2022). Longitudinal changes in mental health and the COVID-19 pandemic: Evidence from the UK Household Longitudinal Study. *Psychological Medicine*, *52*(13), 2549–2558. https://doi.org/10.1017/S0033291720004432

Ellwardt, L., & Präg, P. (2021). Heterogeneous mental health development during the COVID-19 pandemic in the United Kingdom. *Scientific Reports*, *11*(1), Article 15958. https://doi.org/10.1038/s41598-021-95490-w

Fernandes, H. M., & Vasconcelos-Raposo, J. (2013). Factorial validity and invariance of the GHQ-12 among clinical and nonclinical samples. *Assessment*, *20*(2), 219–229. https://doi.org/10.1177/1073191112465768

Gnambs, T., & Staufenbiel, T. (2018). The structure of the General Health Questionnaire (GHQ-12): Two meta-analytic factor analyses. *Health Psychology Review*, *12*(2), 179–194. https://doi.org/10.1080/17437199.2018.1426484

Goldberg, D. P., Gater, R., Sartorius, N., Ustun, T. B., Piccinelli, M., Gureje, O., & Rutter, C. (1997). The validity of two versions of the GHQ in the WHO study of mental illness in general health care. *Psychological Medicine*, *27*(1), 191–197. https://doi.org/10.1017/S0033291796004242

Goldberg, D. P., & Williams, P. (2000). *General Health Questionnaire (GHQ)*. nferNelson.

Graham, J. W. (2003). Adding missing-data-relevant variables to FIML-based structural equation models. *Structural Equation Modeling*, *10*(1), 80–100. https://doi.org/10.1207/S15328007SEM1001_4

Greenberg, M. T., & Abenavoli, R. (2017). Universal interventions: Fully exploring their impacts and potential to produce population-level impacts. *Journal of Research on Educational Effectiveness*, *10*(1), 40–67. https://doi.org/10.1080/19345747.2016.1246632

Holmes, E. A., O'Connor, R. C., Perry, V. H., Tracey, I., Wessely, S., Arseneault, L., Ballard, C., Christensen, H., Cohen Silver, R., Everall, I., Ford, T., John, A., Kabir, T., King, K., Madan, I., Michie, S., Przybylski, A. K., Shafran, R., Sweeney, A., … Bullmore, E. (2020). Multidisciplinary research priorities for the COVID-19 pandemic: A call for action for mental health science. *The Lancet: Psychiatry*, *7*(6), 547–560. https://doi.org/10.1016/S2215-0366(20)30168-1

Hu, L., & Bentler, P. M. (1999). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*, 424–453. https://doi.org/10.1037/1082-989X.3.4.424

Karataş, Z., & Tagay, Ö. (2021). The relationships between resilience of the adults affected by the covid pandemic in Turkey and Covid-19 fear, meaning in life, life satisfaction, intolerance of uncertainty and hope. *Personality and Individual Differences*, *172*, Article 110592. https://doi.org/10.1016/j.paid.2020.110592

Liu, Y., Millsap, R. E., West, S. G., Tein, J.-Y., Tanaka, R., & Grimm, K. J. (2017). Testing measurement invariance in longitudinal data with ordered-categorical measures. *Psychological Methods*, *22*(3), 486–506. https://doi.org/10.1037/met0000075

Lombardo, P., Jones, W., Wang, L., Shen, X., & Goldner, E. M. (2018). The fundamental association between mental health and life satisfaction: Results from successive waves of a Canadian national survey. *BMC Public Health*, *18*(1), Article 342. https://doi.org/10.1186/s12889-018-5235-x

Magson, N. R., Freeman, J. Y. A., Rapee, R. M., Richardson, C. E., Oar, E. L., & Fardouly, J. (2021). Risk and protective factors for prospective changes in adolescent mental health during the COVID-19 pandemic. *Journal of Youth and Adolescence*, *50*(1), 44–57. https://doi.org/10.1007/s10964-020-01332-9

Mäkikangas, A., Feldt, T., Kinnunen, U., Tolvanen, A., Kinnunen, M.-L., & Pulkkinen, L. (2006). The factor structure and factorial invariance of the 12-item General Health Questionnaire (GHQ-12) across time: Evidence from two community-based samples. *Psychological Assessment*, *18*(4), 444–451. https://doi.org/10.1037/1040-3590.18.4.444

Muthén, B. O. (1998). *Mplus technical appendices*.

Neufeld, S. A. S., Clair, M. S., Brodbeck, J., Wilkinson, P., Goodyer, I., & Jones, P. B. (2022). *Measurement invariance in longitudinal bifactor models: Review and application based on the p factor*. https://doi.org/10.17605/OSF.IO/FBD3H

Olino, T. M., Case, J. A. C., Hawes, M. T., Szenczy, A., Nelson, B., & Klein, D. N. (2022). Testing invariance of measures of internalizing symptoms before and after a major life stressor: The impact of COVID-19 in an adolescent and young adult sample. *Assessment*, *29*(7), 1371–1380. https://doi.org/10.1177/10731911211015315

Patel, K., Robertson, E., Kwong, A. S. F., Griffith, G. J., Willan, K., Green, M. J., Di Gessa, G., Huggins, C. F., McElroy, E., Thompson, E. J., Maddock, J., Niedzwiedz, C. L., Henderson, M., Richards, M., Steptoe, A., Ploubidis, G. B., Moltrecht, B., Booth, C., Fitzsimons, E., … Katikireddi, S. V. (2022). Psychological distress before and during the COVID-19 pandemic among adults in the United Kingdom based on coordinated analyses of 11 longitudinal studies. *JAMA Network Open*, *5*(4), Article e227629. https://doi.org/10.1001/jamanetworkopen.2022.7629

Pierce, M., Hope, H., Ford, T., Hatch, S., Hotopf, M., John, A., Kontopantelis, E., Webb, R., Wessely, S., McManus, S., & Abel, K. M. (2020). Mental health before and during the COVID-19 pandemic: A longitudinal probability sample survey of the UK population. *The Lancet: Psychiatry*, *7*(10), 883–892. https://doi.org/10.1016/S2215-0366(20)30308-4

Pierce, M., McManus, S., Hope, H., Hotopf, M., Ford, T., Hatch, S. L., John, A., Kontopantelis, E., Webb, R. T., Wessely, S., & Abel, K. M. (2021). Mental health responses to the COVID-19 pandemic: A latent class trajectory analysis using longitudinal UK data. *The Lancet: Psychiatry*, *8*(7), 610–619. https://doi.org/10.1016/S2215-0366(21)00151-6

Prenoveau, J. M. (2016). Specifying and interpreting latent state–trait models with autoregression: An illustration. *Structural Equation Modeling*, *23*(5), 731–749. https://doi.org/10.1080/10705511.2016.1186550

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Romppel, M., Hinz, A., Finck, C., Young, J., Brähler, E., & Glaesmer, H. (2017). Cross-cultural measurement invariance of the General Health Questionnaire-12 in a German and a Colombian population sample. *International Journal of Methods in Psychiatric Research*, *26*(4), Article e1532. https://doi.org/10.1002/mpr.1532

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Ruiz, F. J., García-Beltrán, D. M., & Suárez-Falcón, J. C. (2017). General Health Questionnaire-12 validity in Colombia and factorial equivalence between clinical and nonclinical participants. *Psychiatry Research*, *256*, 53–58. https://doi.org/10.1016/j.psychres.2017.06.020

Salk, R. H., Hyde, J. S., & Abramson, L. Y. (2017). Gender differences in depression in representative national samples: Meta-analyses of diagnoses and symptoms. *Psychological Bulletin*, *143*(8), 783–822. https://doi.org/10.1037/bul0000102

Sen-Crowe, B., McKenney, M., & Elkbuli, A. (2020). Social distancing during the COVID-19 pandemic: Staying home save lives. *The American Journal of Emergency Medicine*, *38*(7), 1519–1520. https://doi.org/10.1016/j.ajem.2020.03.063

Shevlin, M., & Adamson, G. (2005). Alternative factor models and factorial invariance of the GHQ-12: A large sample analysis using confirmatory factor analysis. *Psychological Assessment*, *17*(2), 231–236. https://doi.org/10.1037/1040-3590.17.2.231

Shevlin, M., Butter, S., McBride, O., Murphy, J., Gibson-Miller, J., Hartman, T. K., Levita, L., Mason, L., Martinez, A. P., McKay, R., Stocks, T. V., Bennett, K. M., Hyland, P., Vallieres, F., Valiente, C., Vazquez, C., Contreras, A., Peinado, V., Trucharte, A., … Bentall, R. P. (2022). Measurement invariance of the Patient Health Questionnaire (PHQ-9) and Generalized Anxiety Disorder scale (GAD-7) across four European countries during the COVID-19 pandemic. *BMC Psychiatry*, *22*(1), Article 154. https://doi.org/10.1186/s12888-022-03787-5

University of Essex, Institute for Social and Economic Research. (2019). *NatCen Social Research, Kantar Public Understanding Society: Waves 1–9, 2009–2018 and Harmonised BHPS: Waves 1–18, 1991–2009* [data collection]. 17th Edition. UK Data Service. SN: 6614. https://doi.org/10.5255/UKDA-SN-6614-18

Xiong, J., Lipsitz, O., Nasri, F., Lui, L. M. W., Gill, H., Phan, L., Chen-Li, D., Iacobucci, M., Ho, R., Majeed, A., & McIntyre, R. S. (2020). Impact of COVID-19 pandemic on mental health in the general population: A systematic review. *Journal of Affective Disorders*, *277*, 55–64. https://doi.org/10.1016/j.jad.2020.08.001

Received October 21, 2022
Revision received March 28, 2023
Accepted April 6, 2023 ■