

Original Article

Cite this article: Breslau J, McBain R, Wong EC, Roth E, Burnam MA, Cefalu MS, Collins RL (2020). A randomised test of the effect of medical v. lay idiom on assessment of perceived mental health condition in the USA. *Epidemiology and Psychiatric Sciences* **29**, e172, 1–8. <https://doi.org/10.1017/S2045796020000840>

Received: 10 June 2020

Revised: 27 August 2020

Accepted: 2 September 2020


Key words:

Epidemiology; mental health; methodological experiment; idiom of distress

Author for correspondence:

Joshua Breslau, E-mail: jbreslau@rand.org

A randomised test of the effect of medical v. lay idiom on assessment of perceived mental health condition in the USA

J. Breslau¹ , R. McBain², E. C. Wong³, E. Roth³, M. A. Burnam³, M. S. Cefalu³ and R. L. Collins³

¹RAND Corporation, Pittsburgh, Pennsylvania, USA; ²RAND Corporation, Boston, Massachusetts, USA and ³RAND Corporation, Santa Monica, California, USA

Abstract

Aims. To test the impact of using different idioms in epidemiological interviews on the prevalence and correlates of poor mental health and mental health service use.

Methods. We conducted a randomised methodological experiment in a nationally representative sample of the US adult population, comparing a lay idiom, which asked about ‘problems with your emotions or nerves’ with a more medical idiom, which asked about ‘problems with your mental health’. Differences across study arms in the associations of endorsement of problems with the Kessler-6 (a validated assessment of psychological distress), demographic characteristics, self-rated health and mental health service use were examined.

Results. Respondents were about half as likely to endorse a problem when asked with the more medical idiom (18.1%) than when asked with the lay idiom (35.1%). The medical idiom had a significantly larger area under the ROC curve when compared against a validated measure of psychological distress than the lay idiom (0.91 v. 0.87, $p = 0.012$). The proportion of the population who endorsed a problem but did not receive treatment in the past year was less than half as large for the medical idiom (7.90%) than for the lay idiom (20.94%). Endorsement of problems differed in its associations with age, sex, race/ethnicity and self-rated health depending on the question idiom. For instance, the odds of endorsing problems were threefold higher in the youngest than the oldest age group when the medical idiom was used (OR = 3.07; 95% CI 1.47–6.41) but did not differ across age groups when the lay idiom was used (OR = 0.76; 95% CI 0.43–1.36).

Conclusion. Choice of idiom in epidemiological questionnaires can affect the apparent correlates of poor mental health and service use. Cultural change within populations over time may require changes in instrument wording to maintain consistency in epidemiological measurement of psychiatric conditions.

Introduction

Interview assessments are critical to psychiatric epidemiology, given the lack of biological measurements of mental health status and the low portion of disorder cases that are seen in clinical settings (Robins, 1985). The design of these instruments is challenging because respondents use a variety of idioms to describe their conditions and may be reluctant to acknowledge or disclose symptoms due to associated stigma (Corrigan *et al.*, 2014; Lewis-Fernandez and Kirmayer, 2019). Some of the terminology used in psychiatric instruments draws on medical terminology by referring explicitly to mental health or psychiatric symptoms. However, medical terminology may be negatively valued or, when it fails to match respondents’ conceptualisations of mental disorder, poorly understood. Idioms of distress also vary across cultures and over time within cultures (Kohrt *et al.*, 2013). To correct for potential limitations of medical terminology, instrument designers use lay idioms, which are more commonly understood in the general population, broader in reference and less weighted with negative connotations (Nichter, 2010). Since lay idioms are selected to be acceptable to respondents their use is likely to result in higher estimates of the prevalence of mental health conditions than medical idioms. In practice, both types of idioms are frequently used together. For instance, the Alcohol Use Disorder and Associated Disabilities Interview Schedule (AUDADIS) module used to assess major depressive episodes combines the lay terms ‘sad’, ‘blue’ and ‘down’ with the more clinical term ‘depressed’ in a single question (Ruan *et al.*, 2008; Grant *et al.*, 2003).

While it is reasonable to expect that greater reliance on lay relative to medical idioms would result in the higher endorsement of mental health conditions, it is not known whether responses to these alternatives differ across population groups. If responses do not differ, then the higher prevalence produced by instruments that rely more on lay idioms may be

© The Author(s), 2020. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives licence (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is unaltered and is properly cited. The written permission of Cambridge University Press must be obtained for commercial re-use or in order to create a derivative work.

preferable, since they will have higher power to detect risk factor associations. However, if responses differ, decisions about idioms may affect identification of risk factors; studies that employ different idioms may identify different risk factors due to language choice rather than actual differences in psychiatric morbidity. For instance, there is evidence that item wording might affect associations between race/ethnicity and perceived unmet need for treatment (Breslau *et al.*, 2018). Changes in response to question-wording may also occur over time, reflecting the change in attitudes towards mental illness (Pescosolido *et al.*, 2010), and thereby influence the assessment of trends in morbidity. The lay idioms of today's instruments are medical idioms from earlier historical periods.

To examine the impact of alternative strategies for inquiring about mental health conditions, we conducted a randomised methodological experiment in the context of a survey of a nationally representative sample of the US adult population. While there is a rich and growing body of experiments focused on questionnaire wording and design in a cross-cultural context, no experimental studies have been conducted of idiom with respect to mental health conditions. In this study, respondents were randomised to receive one of two items, one using a more medical idiom of 'mental health problems' and one using a common lay idiom of 'emotions and nerves'. We then examined differences in the prevalence of endorsement of the two items, their correspondence with a standard assessment of serious psychological distress and with use of mental health services, and their association with common correlates of psychiatric disorders.

Methods

The RAND American Life Panel (ALP) is a probability sample-based panel of about 6000 adult Americans. Panel members respond to surveys via the internet on mobile devices or via computer. Respondents are contacted via email to fill out questionnaires and they are paid quarterly for their participation. A total of 3932 ALP respondents were contacted in February 2019 to respond to the Omnibus survey. Responses were obtained from 2555 respondents, for a response rate of 64.9%. One respondent with missing data on all outcome measures was excluded from the analysis.

Assessments

Perceived mental health

Respondents were randomised to one of two questions regarding their mental health. The questions were designed to tap different strategies that are combined in the instrument used in the National Survey of Drug Use and Health (NSDUH) (Substance Abuse and Mental Health Data Archive, 2018) and the World Mental Health version of the Composite International Diagnostic Instrument (Kessler and Üstün, 2004). Those instruments produce some of the most often-cited estimates of perceived mental health, asking about 'problems with your emotions, nerves, or mental health'. In this study, we separated the more medical term 'mental health' from the lay terms 'emotions and nerves'. Half of respondents were randomised to receive the item: 'Was there ever a time during the past 12 months when you had problems with your mental health?' The other group received the item: 'Was there ever a time during the past 12 months when you had problems with your emotions

or nerves?' We refer to these two items hereafter as MHP and EMO, respectively.

Individual characteristics and self-rated health

Demographic information collected on respondents included gender, age, race/ethnicity, marital status, educational attainment, household income and employment status. Self-rated health (Schnittker and Bacak, 2014) was assessed with a single item and classified as poor/fair, good, or very good/excellent.

Psychological distress

Respondents were administered the Kessler-6 (K6) with reference to the worst month of the past-year. The K6 is a scale of non-specific psychological distress designed to identify cases of clinically significant distress (Kessler *et al.*, 2003). K6 scores, which range from 0 to 24, were used continuously in a receiver operator characteristic (ROC) curve analysis and also classified into three levels, as suggested by prior research (Furukawa *et al.*, 2008) in cross-tabs and regression models: low distress (0–7), mild to moderate distress (8–12) and serious distress (13–24).

Service use

Respondents in both study arms were asked the same question about mental health service use in the past year. This question, which was administered after the question about the mental health condition, referenced both problems with mental health and problems with emotions or nerves: 'In the past 12 months, did you see a professional, such as a physician, counselor, psychiatrist, or social worker for problems with your emotions, nerves, or mental health?' This wording mirrors that used by measures of mental health treatment use in the NSDUH. Service use in the past-year was coded as a dichotomous variable.

Statistical analysis

All analyses were weighted for non-response and to match demographic characteristics of the US population as described in the ALP technical documentation (Pollard and Baird, 2017). Statistical tests were adjusted for the complex sampling design. Differences in the relationship of each item with psychological distress, measured with the K6 as a score ranging from 0 to 24, were examined using ROC curve analysis. A chi-square test was used to test the difference in the area under the ROC curve (AUC) between the two questions. Correspondence between the questions and the K6 was also examined with standard test performance measures (sensitivity, specificity, positive and negative predictive value, and accuracy) at the two K6 thresholds.

Relationships of question endorsement with each of a set of candidate predictors: demographic characteristics (gender, age, marital status, race/ethnicity, educational attainment, employment status, household income), self-rated health and psychological distress, were examined using multivariable logistic regression. Models were estimated in the entire sample, i.e. study arms were combined, with the endorsement of poor mental health (according to the EMO or MHP question) as the outcome. For each candidate predictor, we estimated a model with the candidate predictor, an indicator for question type (whether the respondent was randomised/responding to EMO or MHP), an interaction between the predictor and question type, and main effects of other predictors. These models were used to estimate within-arm odds ratios showing associations between the

candidate predictors and endorsement of each question type and to test for variation in these associations across question type.

Results

The sample was balanced across the two study arms with respect to demographic characteristics, self-rated health and psychological distress (Table 1). Respondents were about half as likely to endorse MHP (18.1%) as EMO (35.1%) and the items were substantially different on most performance metrics (Table 2). Correspondence of endorsement with clinically significant psychological distress, measured by the continuous K6, was stronger for MHP than for EMO. The area under the ROC curve was significantly larger for MHP than for EMO (0.91 *v.* 0.87, $p = 0.012$). Sensitivity is moderate for both items at the lower threshold. MHP reached its maximum accuracy of 0.89 at a threshold of $K6 > 14$. In contrast, EMO reached its maximum accuracy of 0.80 at a much lower threshold of $K6 > 6$. When compared with the commonly used mild/moderate distress cutpoint of $K6 > 7$, MHP had lower sensitivity, higher specificity, higher positive predictive value, lower negative predictive value and higher accuracy than EMO. The same pattern was observed at the $K6 > 12$ (serious distress) threshold. Notably, at the higher K6 threshold, EMO had a positive predictive value of 0.23, indicating that an individual endorsing the EMO question had only a 23% probability of having serious psychological distress as assessed by the K6.

Figure 1 shows how the endorsement of a problem according to each question is related to the use of mental health services. It is notable that two proportions that have important public health implications have quite different magnitudes in the two study arms. First, the proportion of people who endorse a problem and do not use services is 7.90% of the population in the MHP arm and 20.94% in the EMO arm. Second, the proportion of people who do not endorse a problem and nonetheless use mental health services is 6.48% in the MHP arm and 2.59% in the EMO arm.

The endorsement was higher for EMO than MHP at all levels of all demographic characteristics, self-rated health and psychological distress (Table 3). However, variation in the relationship between predictors and endorsement was observed in adjusted regression models for four predictors. First, the female gender is associated with a higher endorsement of EMO, but there is no significant association between gender and MHP. Second, younger age is associated with higher endorsement of MHP, but there is no significant relationship between age and EMO. Third, the relationship between race/ethnicity and endorsement is statistically significant for EMO but not for MHP. Examination of the odds ratios suggests that the difference between the arms is largest for the Non-Hispanic (NH)-Black *v.* NH-White comparison; for EMO the OR = 0.41 (95% CI 0.22–0.77) while for MHP the OR = 0.75 (95% CI 0.22, 2.55). Fourth, self-rated health is associated with both EMO and MHP, but in different ways. Endorsement of MHP is lower for those with ‘good’ health, the middle response category, than those with ‘poor/fair’ health, the lower response categories. However, endorsement of EMO is monotonically related to self-rated health; lower health rating is associated with significantly higher endorsement of EMO. Statistical interactions between predictors and question type are statistically significant for age ($p = 0.004$) and self-rated health ($p = 0.031$), but not for gender ($p = 0.211$) or race/ethnicity ($p = 0.832$).

Table 1. Characteristics of the sample by Study Arm

	MHP		EMO	
	<i>n</i>	%	<i>n</i>	%
Gender				
Male	568	49.4	534	46.8
Female	718	50.6	734	53.2
Age				
20–39	171	31.9	194	32.9
40–59	496	37.7	452	34.6
60 and up	619	30.4	622	32.5
Marital status				
Married/Co-Hab	773	59.6	745	59.2
Sep/Div/Widowed	317	17.3	324	19.3
Never married	196	23.1	199	21.5
Race/ethnicity				
Hispanic	187	24.3	186	23.7
NH-White	912	59.2	895	59.0
NH-Black	121	11.6	127	11.7
NH-Other	66	5.0	60	5.5
Education				
High school or less	186	39.8	172	37.6
Some college	445	25.5	450	28.4
Bachelor's degree or higher	655	34.8	646	34.1
Employment status				
Employed	748	65.4	709	62.7
Retired	354	18.2	386	21.3
Other employment status	184	16.5	173	16.0
Household income				
Less than 35 K	311	27.5	324	27.8
35K–60K	302	21.1	304	22.0
60K–99K	299	24.5	287	23.4
100K+	371	26.9	352	26.8
Self-reported health				
Poor/Fair	184	14.2	181	13.5
Good	423	35.3	405	34.4
Very Good/Excellent	670	50.4	675	52.1
Psychological distress (K6)				
0–7: Low distress	1029	74.5	985	75.9
8–12: Moderate distress	140	12.4	173	15.1
13+: High distress	117	13.0	110	9.1

MHP, Mental health problems; EMO, problems with emotions or nerves. *ns* are unweighted. Percentages are weighted. K6, Kessler-6.

Discussion

Instruments for assessing psychiatric disorders in the general population are critical for informing research and policy,

Table 2. Relationships between endorsement and clinically significant psychological distress (K6) by question idiom

Question Idiom	Endorsement (%)	AUC	Max ACC	K6 at Max ACC	Cutoff at mild/moderate or higher (K6 > 7)					Cutoff at serious (K6 > 12)				
					Sens	Spec	PPV	NPV	ACC	Sens	Spec	PPV	NPV	ACC
MHP	18.1	0.91*	0.89	14	0.57	0.95	0.80	0.87	0.86	0.74	0.90	0.53	0.96	0.88
EMO	35.1	0.87*	0.80	6	0.75	0.78	0.52	0.91	0.77	0.87	0.70	0.23	0.98	0.72

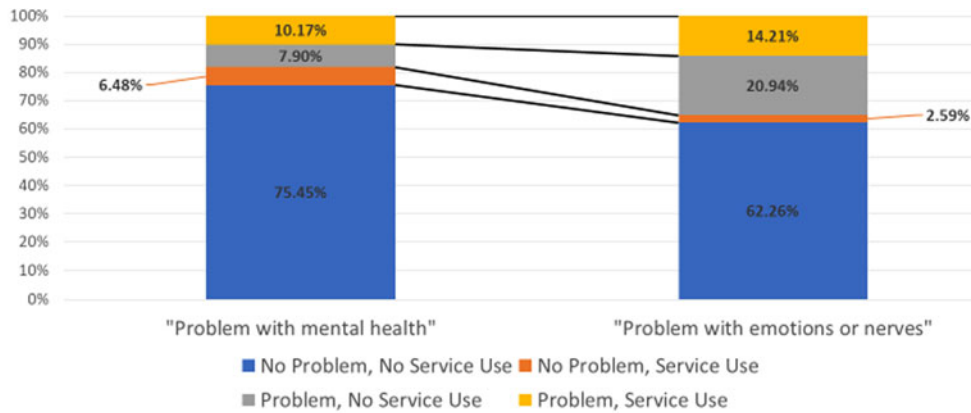
MHP, mental health problems; EMO, emotions or nerves; AUC, area under the ROC curve; ACC, accuracy; K6, Kessler-6 score; Sens, sensitivity; Spec, specificity; PPV, positive predictive value; NPV, negative predictive value.
 *Difference in AUCs is statistically significant at $p = 0.05$ ($\chi^2_{(1)} = 6.313, p = 0.012$).

particularly due to the low rate of treatment and the consequent need to assess community samples to estimate the prevalence and identify risk factors. One of the many challenges in designing such assessments is matching the language of questions to respondents' interpretations and responses in the context of culturally variable and historically shifting practices related to mental health conditions and the language commonly used to describe them. Studies of idioms used to assess mental health problems have largely been limited to cross-cultural settings (Lewis-Fernández et al., 2010), but the same issues are likely to arise within a population as well. This randomised methodological experiment of two idioms for assessing mental health conditions provides some insight that can inform future instrument design strategies as well as interpretation of current evidence.

It is important to note that the two alternatives tested here are not meant to represent mutually exclusive alternatives or as stand-alone assessments of mental health conditions. Rather, this experiment was designed to test differences in responses to the underlying idioms that are in practice commonly used together, particularly in longer diagnostic instruments such as the WMH-CIDI (Kessler and Üstün, 2004) or the AUDADIS (Ruan et al., 2008). Used alone, these are not a valid assessment of mental health status. Our findings should be interpreted as indicating that, other things being equal, an instrument using a more medical idiom would be likely to produce results that differ from one using a more lay idiom in the ways that responses to the MHP and EMO questions differ from each other. As expected, the MHP item more closely corresponds with our measure of clinically significant psychological distress, a measure that is validated by a medical diagnosis. However, we do not intend to suggest that either approach is inherently better than the other. The selection of idiom should be determined by the purpose of the item within a questionnaire and the overall purpose of the study.

The more medical idiom produced not only a much lower prevalence of endorsement of poor mental health, as expected but also a closer correspondence with a widely used and validated measure of clinically significant psychological distress. The lower test performance of EMO relative to MHP is particularly striking at higher levels of distress, where clinical significance is most challenging to establish and results are most relevant to public health. In addition, service use was more common among people who endorsed MHP than among those who endorsed EMO (although the prevalence of service use did not differ across study arms). This finding, based on actual help-seeking behavior rather than a survey instrument, provides additional external validation of the difference between the items. The relationship between endorsement and service use has public health importance. Although we do not interpret endorsement as a valid indicator of the need for treatment, the findings suggest that an instrument with a more lay idiom would result in a larger estimate of unmet need for treatment than a similar instrument that employed a more medical idiom.

The results also suggest that varying the question idiom impacts the apparent distribution of mental health conditions across demographic groups. There is evidence of differential response with respect to gender, age, race/ethnicity and self-rated health. It is important to note that the evidence of differences related to gender and race/ethnicity is relatively weak; differences were observed in significance between the two arms, but the statistical interaction did not reach significance. In both cases, the ORs are in the same direction for both arms of the study, but there is a difference in magnitude and statistical significance.



Results from ALP survey (n=2,555). Respondents randomized to question idiom.

Fig. 1. Question idiom and service use. Results from ALP survey (n = 2555). Respondents randomised to question idiom.

These findings are important to note because they could influence the findings of epidemiological studies in specific ways. The findings suggest that an instrument with a more lay idiom would find larger differences between males and females and between NH-Blacks and NH-Whites than an instrument with a more medical idiom. The more lay idiom would find that females have a relatively higher prevalence than males and that NH-Blacks have relatively lower prevalence relative to NH-Whites. These potential measurement effects are of interest because both gender (Riecher-Rössler, 2017) and race/ethnicity (Breslau *et al.*, 2005) are well-established risk factors for a broad range of psychopathology.

The evidence for effects of idiom on associations of age and self-rated health with mental health problems is stronger. For these two factors, we find statistically significant interactions in addition to qualitatively meaningful patterns of association. With respect to age, the medical idiom results in a 2-to-3-fold increase in the apparent prevalence of mental health conditions between the oldest and youngest age groups while the lay idiom results did not indicate a significant difference across the same age groups. This finding suggests a dramatic change across birth cohorts in response to different idioms for mental health conditions; younger cohorts are much more likely than older cohorts to endorse medical idioms for mental health conditions. This may signal a decrease in mental health stigma across cohorts, as other studies have reported (Lipson *et al.*, 2019). It is also notable that the medical idiom has not supplanted or replaced the lay idiom but added to it, i.e. there is no decrease in endorsement of EMO concurrent with the increase in endorsement of MHP. Large birth cohort differences in the prevalence of psychiatric disorders have been a consistent finding of epidemiological surveys and may be impacted by these historical trends (Kessler *et al.*, 2005). Changing responses to common survey items may play a role in apparent historical increases in mental health problems among younger cohorts.

The relationship between self-rated health and endorsement of a problem varies in shape between the two question idioms. The more medical idiom results in a 'u-shape' relationship, where reporting 'good' self-rated health is associated with lower odds of endorsing a mental health condition than those with 'poor' or 'fair' self-rated health and those with 'very good' or 'excellent'

self-rated health, although this last difference does not reach statistical significance. In contrast, there is a strong monotonic inverse relationship between self-rated health and endorsement of the lay idiom for mental health conditions. This finding suggests that more medical idiom leads respondents to more clearly differentiate mental health from physical health conditions. Moreover, instruments that use different idioms could produce different results regarding mental health status or mental/physical health comorbidity due to the physical health status of the respondents.

Findings from this study should be interpreted in light of its limitations. First, as noted above the tested alternatives are not meant to be used in isolation to assess mental health status or perceived need for treatment, and they do not represent the full range of potential idioms or combinations of idioms that can be used to construct items and assessment instruments. However, it will never be practical to test all the possible alternative measurement strategies against one another. The contrast between MHP and EMO can inform the design process, by furnishing information on how these idioms affect respondent responses, but the results should not be interpreted as favouring one of these items over the other. Undoubtedly, both strategies will continue to be broadly used.

Second, we assessed the performance of the two items against the K6, as if the K6 represented a 'gold standard' clinical evaluation. While the use of the same validated standard to assess both questions is a strength, as a standard, the K6 has known limitations. The K6 does not cover the full range of psychopathology and it does not address functional impairments (Kessler *et al.*, 2010). It is possible that the two idioms tested here would vary in their relative performance across measures of impairment, different types of psychiatric disorder, or indicators of well-being such as life satisfaction and purpose. We were unable to explore such variation in this study.

Third, the low response rate lowers the generalisability of the results to the US population. The ALP panel is recruited using population sampling methods and non-response weights are used to match the sample distribution to that of the US general population with respect to gender, age, ethnicity, education, household income and household size. Low response rates for some groups would lower our power to detect within-group

Table 3. Predictors of endorsement of mental health problems and emotional problems

	Endorsement		Adjusted odds ratios					
	MHP	EMO	MHP			EMO		
	%	%	AOR	95% CI	<i>p</i> -value*	AOR	95% CI	<i>p</i> -value*
Gender								
Female	18.5	42.8	1.15	(0.63, 2.08)	0.655	1.87	(1.16, 3.02)	0.010
Male	17.6	26.5	Reference			Reference		
Age								
20–39	28.3	36.4	3.07	(1.47, 6.41)	0.003	0.76	(0.43, 1.36)	0.650
40–59	19.3	38.4	2.35	(1.30, 4.25)		0.95	(0.58, 1.55)	
60 and up	5.8	30.5	Reference			Reference		
Marital status								
Married/Co-Hab	13.7	35.3	0.51	(0.23, 1.16)	0.258	1.52	(0.75, 3.10)	0.285
Sep/Div/Widowed	18.6	35.4	0.68	(0.31, 1.50)		1.11	(0.51, 2.41)	
Never married	28.9	34.6	Reference			Reference		
Race/ethnicity								
Hispanic	24.6	38.8	1.02	(0.50, 2.08)	0.764	0.82	(0.40, 1.70)	0.022
NH-Black	17.0	31.1	0.75	(0.22, 2.55)		0.41	(0.22, .77)	
NH-Other	9.7	24.0	0.57	(0.18, 1.75)		0.47	(0.22, 1.03)	
NH-White	16.3	35.5	Reference			Reference		
Education								
BA or higher	19.3	33.5	2.22	(0.94, 5.22)	0.187	1.27	(0.72, 2.23)	0.621
Some college	21.2	37.1	1.75	(0.77, 3.99)		1.04	(0.59, 1.85)	
High school or less	15.0	35.1	Reference			Reference		
Employment status								
Retired	6.8	27.5	0.57	(0.27, 1.22)	0.317	1.24	(0.74, 2.06)	0.697
Other	22.8	51.0	1.01	(0.57, 1.81)		0.88	(0.45, 1.71)	
Employed	20.0	33.7	Reference			Reference		
Household income								
Less than 35K	22.1	36.7	0.47	(0.19, 1.16)	0.309	0.42	(0.20, .85)	0.092
35K–60K	17.9	36.5	0.59	(0.23, 1.51)		0.82	(0.42, 1.61)	
60K–99K	12.2	35.8	0.49	(0.21, 1.10)		0.86	(0.49, 1.51)	
100K+	19.5	32.0	Reference			Reference		
Self-reported health								
Poor/Fair	39.7	60.4	1.83	(0.88, 3.80)	0.010	3.05	(1.55, 5.98)	0.001
Good	16.4	43.6	0.63	(0.30, 1.30)		2.07	(1.26, 3.40)	
Very Good/Excellent	13.1	23.6	Reference			Reference		
K6 distress								
13+: High distress	73.6	87.2	48.35	(20.17, 115.90)	<0.001	17.36	(7.95, 37.92)	<0.001
8–12: Moderate distress	39.6	68.3	10.66	(5.09, 22.33)		6.23	(3.45, 11.25)	
0–7: Low distress	4.8	22.3	Reference			Reference		

MHP, mental health problems; EMO, problems with emotions or nerves; AOR, adjusted odds ratio. Percentages are weighted. K6, Kessler-6.

**P*-values correspond to Wald Chi-Square tests for significant associations between each variable and item endorsement within each study arm. These tests are conducted in survey logistic regression models with interactions between the variable of interest and study arm along with adjustment for other variables in the table.

differences. In addition, the study was limited to an English language instrument, which meant the exclusion of respondents without proficiency in English. Unfortunately, this exclusion prevents us from examining difference related to nativity and language, where we would expect to find the largest differences in response to different question idioms. Differences related to the nativity in psychiatric morbidity (Breslau *et al.*, 2009) and perception of the need for mental health treatment (Breslau *et al.*, 2017) are larger than differences related to race/ethnicity among the US born population. Moreover, the findings clearly cannot be directly generalised to other settings where the functioning of various medical and lay idioms is different from the USA.

Historically, research on idioms used to express psychiatric conditions has largely been qualitative in nature (Nichter, 2010), though the volume of research using standardised measures has increased in recent years (Kohrt *et al.*, 2014). Standardised measures have also been used to explore lay theories of mental illness (Furnham and Telford, 2012). Qualitative methods have been critical in identifying and interpreting cultural models of illness but limited in population sampling and analysis of clinical significance. Future qualitative research could contribute greater depth to our understanding of how the use of different idioms varies across the population and is connected to more elaborate cultural models of health and illness. Advances are most likely to come through iteration over time between qualitative explications of idioms of distress and quantitative approaches that examine population patterns.

The challenge of designing instruments to assess mental health conditions and individuals' perceptions of them in the general population is an ongoing and shifting one. The idioms in which poor mental health is expressed vary culturally, but they are not fixed within cultures, and variations in functioning may affect survey results in ways that have important public health implications. Along with clinical calibration studies, studies of responses to different question idioms is an important part of the instrument validation process and helpful for interpreting the results obtained with common measures.

Data. Data used in this study are available through the American Life Panel website: <https://www.rand.org/research/data/alp.html>.

Acknowledgements. None.

Financial support. This research was supported by grants from the National Institute of Minority Health and Health Disparities (R01 MD010274, PI: Breslau) and the National Institute of Mental Health (R01 MH104381, PI: Collins).

Conflict of interest. The authors have no potential conflicts of interest to disclose.

Ethical standards. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

References

- Breslau J, Kendler KS, Su M, Aguilar-Gaxiola S and Kessler RC (2005) Lifetime risk and persistence of psychiatric disorders across ethnic groups in the United States (vol 35, pg 317, 2005). *Psychological Medicine* **35**, 766–766.
- Breslau J, Borges G, Hagar Y, Tancredi D and Gilman S (2009) Immigration to the USA and risk for mood and anxiety disorders: variation by origin and age at immigration. *Psychological Medicine* **39**, 1117–1127.
- Breslau J, Cefalu M, Wong EC, Burnam MA, Hunter GP, Florez KR and Collins RL (2017) Racial/ethnic differences in perception of need for mental health treatment in a US national sample. *Social Psychiatry and Psychiatric Epidemiology* **52**, 929–937.
- Breslau J, Stein BD, Burns RM, Collins RL, Han B, Yu H and Mojtabai R (2018) Examining contradictory evidence on racial/ethnic differences in perceived need for behavioral health treatment. *International Journal of Methods in Psychiatric Research* **27**, e1743.
- Corrigan PW, Druss BG and Perlick DA (2014) The impact of mental illness stigma on seeking and participating in mental health care. *Psychological Science in the Public Interest* **15**, 37–70.
- Furnham A and Telford K (2012) Public attitudes, lay theories and mental health literacy: the understanding of mental health. In L'Abate Luciano (ed.), *Mental Illnesses—Understanding, Prediction and Control*. Croatia: InTech, pp. 3–22.
- Furukawa TA, Kawakami N, Saitoh M, Ono Y, Nakane Y, Nakamura Y, Tachimori H, Iwata N, Uda H, Nakane H, Watanabe M, Naganuma Y, Hata Y, Kobayashi M, Miyake Y, Takeshima T and Kikkawa T (2008) The performance of the Japanese version of the K6 and K10 in the World Mental Health Survey Japan. *International Journal of Methods in Psychiatric Research* **17**, 152–158.
- Grant B, Dawson D, Stinson F, Chou P, Kay W and Pickering R (2003) The alcohol use disorder and associated disabilities interview schedule-IV (AUDADIS-IV): reliability of alcohol consumption, tobacco use, family history of depression and psychiatric diagnostic modules in a general population sample. *Drug and Alcohol Dependence* **71**, 7–16.
- Kessler R and Üstün T (2004) The world mental health (WMH) survey initiative version of the world health organization (WHO) composite international diagnostic interview (CIDI). *International Journal of Methods in Psychiatric Research* **13**, 93–121.
- Kessler RC, Barker PR, Colpe LJ, Epstein JF, Gfroerer JC, Hiripi E, Howes MJ, Normand SLT, Manderscheid RW, Walters EE and Zaslavsky AM (2003) Screening for serious mental illness in the general population. *Archives of General Psychiatry* **60**, 184–189.
- Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR and Walters EE (2005) Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey replication. *Archives of General Psychiatry* **62**, 593–602.
- Kessler RC, Green JG, Gruber MJ, Sampson NA, Bromet E, Cuitan M, Furukawa TA, Gureje O, Hinkov H, Hu CY, Lara C, Lee S, Mneimneh Z, Myer L, Oakley-Browne M, Posada-Villa J, Sagar R, Viana MC and Zaslavsky AM (2010) Screening for serious mental illness in the general population with the K6 screening scale: results from the WHO World Mental Health (WMH) survey initiative. *International Journal of Methods in Psychiatric Research* **19**, 4–22.
- Kohrt BA, Rasmussen A, Kaiser BN, Haroz EE, Maharjan SM, Mutamba BB, de Jong JTVM and Hinton DE (2013) Cultural concepts of distress and psychiatric disorders: literature review and research recommendations for global mental health epidemiology. *International Journal of Epidemiology* **43**, 365–406.
- Kohrt BA, Rasmussen A, Kaiser BN, Haroz EE, Maharjan SM, Mutamba BB, de Jong JTVM and Hinton DE (2014) Cultural concepts of distress and psychiatric disorders: literature review and research recommendations for global mental health epidemiology. *International Journal of Epidemiology* **43**, 365–406.
- Lewis-Fernandez R and Kirmayer LJ (2019) Cultural concepts of distress and psychiatric disorders: understanding symptom experience and expression in context. *Transcultural Psychiatry* **56**, 786–803.
- Lewis-Fernández R, Gorritz M, Raggio GA, Peláez C, Chen H and Guarnaccia PJ (2010) Association of trauma-related disorders and dissociation with four idioms of distress among Latino psychiatric outpatients. *Culture, Medicine, and Psychiatry* **34**, 219–243.
- Lipson SK, Lattie EG and Eisenberg D (2019) Increased rates of mental health service utilization by U.S. College Students: 10-Year population-level trends (2007–2017). *Psychiatric Services* **70**, 60–63.
- Nichter M (2010) Idioms of distress revisited. *Culture, Medicine, and Psychiatry* **34**, 401–416.
- Pescosolido BA, Martin JK, Long JS, Medina TR, Phelan JC and Link BG (2010) A disease like any other? A decade of change in public reactions to

- schizophrenia, depression, and alcohol dependence. *American Journal of Psychiatry* **167**, 1321–1330.
- Pollard M and Baird MD** (2017) *The RAND American Life Panel: Technical Description*. Santa Monica, CA: RAND Corporation.
- Riecher-Rössler A** (2017) Sex and gender differences in mental disorders. *The Lancet Psychiatry* **4**, 8–9.
- Robins LN** (1985) Epidemiology - Reflections on testing the validity of psychiatric interviews. *Archives of General Psychiatry* **42**, 918–924.
- Ruan W, Goldstein R, Chou S, Smith S, Saha T, Pickering R, Dawson D, Huang B, Stinson F and Grant B** (2008) The alcohol use disorder and associated disabilities interview schedule-IV (AUDADIS-IV): reliability of new psychiatric diagnostic modules and risk factors in a general population sample. *Drug and Alcohol Dependence* **92**, 27–36.
- Schnittker J and Bacak V** (2014) The increasing predictive validity of self-rated health. *PLoS One* **9**, e84933.
- Substance Abuse and Mental Health Data Archive** (2018) “National Survey on Drug Use and Health (NSDUH).” Retrieved July 5, 2018, 2018, from <https://www.datafiles.samhsa.gov/study-series/national-survey-drug-use-and-health-nsduh-nid13517>.