Studi sperimentali

Measuring depression with questions about well-being: a study on psychiatric outpatients

Valutazione della depressione attraverso domande sul benessere: uno studio su pazienti psichiatrici

EMANUELE TAROLLA[1], MASSIMO BIONDI[2,3], ELISA FABI[2], ILARIA GAVIANO[2],
ANTONELLA GIGANTESCO[1], LORENZO TARSITANI[2,3], ANGELO PICARDI[1]

E-mail: emanuele.tarolla@gmail.com

[1]Mental Health Unit, Centre of Epidemiology Surveillance and Health Promotion, Italian National Institute of Health, Rome
[2]Department of Neurology and Psychiatry, Sapienza University of Rome
[3]Umberto I Policlinico di Roma Hospital, Rome

SUMMARY. Background. Most depression rating scales are characterized by negatively-phrased questions, exploring the presence of various symptoms. Questions such as those regarding suicidal ideation or painful experiences may reduce acceptability or even lead the reader to withdraw participation in the study. Although positively-worded items may be useful, it should be acknowledged that without formal testing they cannot be assumed to be equivalent to negatively-worded ones. The aim of the present study was to test the reliability and validity of a depression rating scale including only positively-phrased items. Methods. Two groups were enrolled in the study: the first comprised 104 adult psychiatric outpatients, the second 88 undergraduate students. All participants completed the depression scale of the Patient Health Questionnaire, the Zung Self-Rating Depression Scale, and the Positively-phrased Depression Scale (PDS), a 10-item self-report instrument in which the items are phrased in a positive way to reflect the absence of symptoms. Psychiatric outpatients also were rated by their clinician on the Hamilton Depression Rating Scale. Results. The internal consistency of the PDS was satisfactory. The correlations between scores on the PDS and on the other depression scales were moderate to high. Mean PDS scores of patients with a diagnosis of depressive disorder were significantly higher than those of patients with other mental disorders. Conclusions. Despite some limitations, this study suggests that the PDS a valid and reliable instrument which might prove particularly useful for the assessment of depressive symptoms in studies where issues of acceptability are important, such as studies on non-clinical populations, occupational samples, and patients drawn from non-psychiatric settings.

KEY WORDS: depression, rating scales, positive affect.

RIASSUNTO. Introduzione. La maggior parte degli strumenti di valutazione della depressione utilizza domande formulate negativamente che esplorano la presenza dei vari sintomi. Domande come quelle sull'ideazione suicidaria o su esperienze penose possono ridurre l'accettabilità o persino indurre il soggetto a ritirarsi dallo studio. Sebbene domande formulate positivamente potrebbero essere utili, la loro equivalenza a quelle formulate negativamente non può essere assunta senza essere saggiata formalmente. Obiettivo di questo studio è verificare l'affidabilità e la validità di uno strumento di valutazione della depressione costituito da sole domande formulate positivamente. Metodi. Sono stati reclutati due gruppi: il primo costituito da 104 pazienti psichiatrici ambulatoriali, il secondo da 88 studenti universitari. Tutti i partecipanti hanno compilato la scala della depressione del Patient Health Questionnaire, la Zung Self-Rating Depression Scale, e la Positively-phrased Depression Scale (PDS), uno strumento di autovalutazione costituito da 10 item, formulati positivamente in modo da riflettere l'assenza dei vari sintomi. I pazienti sono stati inoltre valutati dal loro psichiatra mediante la Hamilton Depression Rating Scale. Risultati. La coerenza interna della PDS è risultata soddisfacente. Le correlazioni tra i punteggi della PDS e quelli delle altre scale di valutazione della depressione sono risultate moderate o elevate. I punteggi medi alla PDS dei pazienti con una diagnosi di disturbo depressivo sono risultati significativamente maggiori di quelli dei pazienti con altri disturbi mentali. Conclusioni. Tenendo presenti alcune limitazioni, questo studio suggerisce che la PDS sia uno strumento valido e affidabile che potrebbe risultare particolarmente utile per la valutazione della sintomatologia depressiva in situazioni dove l'accettabilità è una questione importante, come negli studi su popolazioni non cliniche, popolazioni del settore lavorativo e pazienti tratti da contesti clinici non psichiatrici.

PAROLE CHIAVE: depressione, scale di valutazione, affettività positiva.

INTRODUCTION

Rating scales are a useful tool to precisely and reliably assess patients' symptoms (1). In psychiatry, systematic use of rating scales allows to monitor clinical response to treatment. Their use along with treatment algorithms into the system of patient care named measurement-based care, is supposed to help to improve depression detection, rationalize decision-making, optimize treatment and improve outcomes (2).

Depression rating scales were developed in the late 1950s with the purpose of measuring the severity of depressive disorders and the changes during drug treatment (3). To date, more than 100 depression rating scale are at the clinician's disposal for diagnostic purpose, to evaluate severity, and to assess changes (4). While there are many ways of classifying scales, one the most important distinctions is between clinician-administered and self-rated scales (5). When compared to clinician-administered instruments, self-rated scales for depression offer some advantages with respect to ease of administration, cost, and time required for the assessment.

Two key issues when using a rating scale are its comprehensibility and acceptability. Therefore, for a self-rating scale it is important to use a clear language, and the scale should not require more than very basic reading skills. Also, terms that might offend or prejudice people should be avoided (6). Most depression rating scales are characterized by negatively-phrased questions, which explore the presence of different symptoms of depressive disorders. Such scales usually include questions about depressed mood and unhappiness, feelings of guilt, pessimistic thoughts, preoccupation with health, and suicidal thoughts. Such questions, especially those regarding suicidal ideation or painful experiences of hopelessness and personal failure, may reduce the acceptability of the assessment instrument in studies carried out on non-clinical populations or patients drawn from non-psychiatric settings. Although to date there is a scarcity of information about the acceptability of the most frequently used rating scales for depression, it is not unlikely that some negatively-phrased questions may cause feelings of distress or embarrassment in the reader who, in turn, may choose to not respond to such questions or even withdraw participation in the study. Also, the responses to these questions may be influenced by social desirability bias and the respondents may minimize symptom severity (7).

Some methodological issues might also arise when choosing between positive and negative phrasing of questions. Although positively-worded items may be useful, it should be acknowledged that without formal testing they cannot be assumed to be equivalent to negatively-worded ones. In a rating scale for depression, the "usually" answer to the positively-worded question "Did you sleep well recently?" is not necessary equivalent to the "never" or "rarely" answer to the negatively-worded question "Did you have difficulty sleeping recently?" (8). In depression rating scales, positive and negative phrasings of items usually refer to the constructs of positive and negative affect, which are considered as the most general dimensions describing affective experience (9). While positive affect reflects the extent to which a person feels enthusiastic, active and alert, negative affect is a general dimension of subjective distress and unpleasurable engagement, encompassing aversive mood states such as anger, contempt, disgust, guilt, fear and nervousness (10). Both high negative and low positive affect seem to be related to depression, whereas only negative affect seems to be related to anxiety (11). It is worth noting that this concept may be not applicable in some populations when positive or negative affect are assessed by specifically-phrased items. For example, a study showed that the Japanese response to the positively worded items of the Center for Epidemiologic Studies Depression Scale (CES-D) markedly differed from those of American workers, as the first group was more likely to choose less positive response alternatives to positive items, while the responses to negatively worded items were generally comparable in the two groups (12). In another study, depressed Japanese patients showed scores similar to those of healthy controls on positive affect items of CES-D, but after rephrasing the positive affect items to negative ones, depressed Japanese patients were found to score higher than the controls (13).

The aim of the present study was to evaluate if it is possible to accurately and validly measure depressive symptoms by using a rating scale composed by positively-phrased items that do not directly refer to psychopathological symptoms.

METHODS

Participants

Two distinct groups were enrolled in the study. Inclusion criteria were age between 18 and 65 years and at least 8 years of education. All participants gave their written informed consent to participate. The first group comprised 104 outpatients attending the psychiatric outpatient facility of the Sapienza – University of Rome between April and May 2010 (58 were females and 46 males, with a mean age of 24.1 ± 14.7 years). The second group included 88 un-

Tarolla E et al.

dergraduate students (54 were females and 34 males, with a mean age of 24.1 ± 3.1 years). More details about the samples are reported in **Tables 1** and **2**.

Assessment instruments

The Positively-phrased Depression Scale (PDS) is an instrument developed to measure depression using positively-phrased questions. It consists of 10 self-report items, each scored on a 5-point scale. The items inquire about the main symptoms of depression, including depressed mood, loss of interest and pleasure, feelings of unhappiness, loss of energy, feelings of uselessness, unrefreshing sleep, difficulty in starting the day, agitation, difficulty concentrating, and retardation. The time covered is the last month. The items are phrased in a positive way to reflect the absence of symptoms (e.g., the question on feelings of agitation was phrased as 'Did you feel calm and relaxed?', while the question on depressed mood was phrased as 'Did you feel happy and in a good mood?', with five possible answers ranging form 'never' to 'most of the time') to reduce the emotional impact of questionnaire completion. The items are scored inversely so that higher scores indicate greater severity of depressive symptoms. In a previous study, test-retest reliability at the item level after two weeks was found to be 0.60 or higher for all items (14).

The 9-item depression scale of the Patient Health Questionnaire (PHQ-9) is a self-report tool aimed to evaluate the presence and severity of the nine DSM-IV-TR criterion symptoms for a major depressive episode, namely anhedonia, depressed mood, trouble sleeping, feeling tired, change in appetite, guilt or worthlessness, trouble concentrating, feeling slowed down or restless, and suicidal thoughts. Each item is rated on a 4-point scale (from 0=Not at all to 3=Nearly every day), with a total score ranging from 0 to 27 (15).

The Zung Self-Rating Depression Scale (ZDS) is a 20-item self-report questionnaire covering affective, psychological and somatic symptoms associated with depression. Each item is scored on a Likert scale ranging from 1 to 4, with a total score ranging from 20 to 80 (16).

The Hamilton Depression Rating Scale (HDRS) is the most commonly used clinician-administered depression scale (17). The scale consists of 17 items which may be scored on a scale from 0 to 4 or 0 to 2. The total score ranges from 0 to 50.

Table 1. Socio-demographic and clinical characteristics of the first sample

		N	%	Mean ± SD
Sex	M	46	*44.0*	
	F	58	*56.0*	
Age				50.5 ± 14.7
Education	None	1	*1.0*	
	Primary school	7	*6.7*	
	Junior high school	28	*26.9*	
	Senior high school	40	*38.5*	
	University degree	21	*20.2*	
	Total	97	*93.3*	
	Missing	7	*6.7*	
	Total	104	*100.0*	
Marital status	Single	31	*29.8*	
	Married	50	*48.1*	
	Separated/divorced	10		
	Widowed	5		
	Total	96	*92.3*	
	Missing	8	*7.7*	
	Total	104	*100.0*	
Diagnosis	No axis I diagnosis	2	*1.9*	
	Schizophrenia	4	*3.8*	
	Other non-affective psychoses	7	*6.7*	
	Depressive disorders	30	*28.8*	
	Anxiety disorders	18	*17.3*	
	Schizoaffective disorder	2	*1.9*	
	Bipolar disorder	14	*13.5*	
	Adjustment disorder	9	*8.7*	
	Other disorders	9	*8.7*	
	Total	95	*91.3*	
	Missing	9	*8.7*	
	Total	104	*100.0*	

Table 2. Socio-demographic characteristics of the second sample

		N	%	Mean ± SD
Sex	M	34	*39.0*	
	F	54	*61.0*	
Age				24.1 ± 3.1
Education	Senior high school	70	*79.5*	
	University degree	9	*10.2*	
	Total	79	*89.9*	
	Missing	9	*10.2*	
	Total	88	*100.0*	
Marital status	Single	79	*89.9*	
	Missing	9	*10.2*	
	Total	88	*100.00*	

Procedure

The participants belonging to the first group completed the PDS, PHQ-9 and ZDS before starting a psychiatric consultation where the HDRS was administred. Participants in the second group completed the PDS, PHQ-9 and ZDS before attending a lesson.

Statistical analysis

All analyses were performed using the SPSS package for Windows, version 17.0.

The internal consistency of PDS was assessed in both samples using coefficient alpha. For every item, item-total correlation and the alpha value by omitting the item were calculated.

The PDS convergent validity was assessed in both groups by calculating Pearson's correlations with the criterion measures, i.e. the PHQ-9 and the ZDS. In the psychiatric outpatient group the HDRS was also used as a further criterion measure. To assess discriminant validity, the Kruskal-Wallis test was used to compare mean PDS scores of patients with depressive disorders with those of patients with other mental disorders.

Table 3. PDS item statistics

Item	First sample		Second sample	
	Item-total correlation	Cronbach's alpha with the item excluded	Item-total correlation	Cronbach's alpha with the item excluded
1	.731	.888	.536	.704
2	.527	.900	.499	.704
3	.689	.890	.466	.710
4	.689	.890	.590	.696
5	.782	.885	.592	.693
6	.677	.891	.191	.753
7	.738	.887	.450	.713
8	.674	.891	.294	.736
9	.513	.901	.207	.751
10	.565	.898	.301	.735

RESULTS

All patients belonging to the clinical sample completed the PDS and the ZDS. One of them did not return a completed PHQ-9. The HDRS was available for 60 patients. All participants belonging to the second group completed the PDS, PHQ-9 and ZDS.

The internal consistency of the PDS was satisfactory, with a coefficient alpha of .90 in the clinical sample, and .74 in the non-clinical sample. The lower value observed in the second sample is likely related to the effect of the lower range of item scores on the inter-item correlations. The results of single-item analysis are shown in **Table 3**.

In the clinical sample, the item-total correlations were moderate to high, while in the non-clinical sample the correlations were moderate on average. Under no circumstances the deletion of a single item would have resulted in a substantial increase in coefficient alpha, as shown in **Table 3**.

With respect to convergent validity, the correlations between scores on the PDS and on the criterion measures were moderate to high. The correlation with the ZDS was .76 in the clinical sample and .66 in the non-clinical sample, respectively. The correlation with the PHQ-9 was .76 in the clinical sample and .64 in the non-clinical sample, respectively. In the clinical sample, the correlation between the PDS and the

HRDS was .55. All the correlations were highly significant (p<.001).

As shown in **Table 4**, in an analysis focused on well-defined diagnostic groups, patients with a diagnosis of depressive disorder scored significantly higher than those with other mental disorders. The items 6, 7, and 10 showed the highest discriminatory power between patient groups (p=.03, p=.08, and p=.04, respectively).

DISCUSSION

Our study provided preliminary evidence in support of the reliability and validity of the PDS as an alternative rating instrument to measure depressive symptoms. First, the moderate to high correlations with the other depression rating scales provided evidence of convergent validity for the PDS. It is noteworthy that the PDS was found to correlate not only with other established self-report rating scales, but also with a clini-

Table 4. PDS mean and standard deviation scores in diagnostic groups (second sample)

Psychiatric diagnoses	Mean (± SD)
Schizophrenia and other non-affective psychoses (n=11)	14.00 (± 11.75)
Depressive disorders (n=30)	19.67 (± 8.08)
Anxiety disorders (n=18)	16.28 (± 8.78)
Other disorders (n=9)	14.25 (± 8.81)
Total (n=68)	17.13 (± 9.14)

Tarolla E et al.

cian-administered scale such as the HDRS. Second, the discriminant validity of the PDS was supported by the finding of a satisfactory discriminatory power not only between healthy subjects and psychiatric patients, but also between patients with depressive disorders and patients with other mental disorders. Also, the reliability of the PDS is supported by the finding of a high coefficient alpha in both samples.

The finding that the correlations between the PDS and the other depression rating scales were substantial, but not redundant, seems to support the notion that, although inversely correlated, psychopathological distress and psychological well-being do not constitute the ends of a unipolar dimension but are at least partially independent. In this regard, the significantly higher PDS scores observed in patients with depressive disorder as compared with patients affected by anxiety disorders seems to corroborate the hypothesis that positive affect rather than negative affect may differentiate depressive from anxiety disorders (18).

It should be acknowledged that the present study has some limitations. Firstly, the reliability of the PDS was evaluated only in terms of internal consistency, and not of stability of scores over time. Future studies should assess the test-retest reliability of the PDS. Second, in the present study the responsiveness (i.e. the sensitivity to change in clinical conditions, a key feature for clinical studies) of the PDS was not assessed. Future research should focus on the sensitivity to change of the PDS in clinical conditions during pharmacological or psychotherapeutic treatment. Third, the narrow age range and the medium to high education level of the nonclinical sample may reduce the generalizability of the results. In the future, the PDS should be tested on a non-clinical population with a broader age and education range.

With these limitations in mind, the results of the present study suggest that the PDS a valid and reliable instrument which, thanks to its distinctive characteristics, might prove particularly useful for the assessment of depressive symptoms in studies where issues of acceptability are important, such as studies on non-clinical populations, occupational samples, and patients drawn from non-psychiatric settings.

REFERENCES

1. Gelenberg AJ. The search for knowledge: developing the American Psychiatric Association's practice guideline for major depressive disorder. J Clin Psychiatry 2008; 69: 1658-9.
2. Gelenberg AJ. Using assessment tools to screen for, diagnose, and treat major depressive disorder in clinical practice. J Clin Psychiatry 2010; 71 Suppl E1: e01.
3. Snaith P. What do depression rating scale measure? Br J Psychiatry 1993; 163: 293-8.
4. Demyttenaere K, De Fruyt J. Getting what you ask for: on the selectivity of depression rating scales. Psychother Psychosom 2003; 72: 61-70.
5. Hamilton M. Comparative value of rating scales. Br J Clin Pharmacol 1976; 1 Suppl 1: 58-60.
6. Keszei AP, Novak M, Streiner DL. Introduction to health measurement scales. J Psychosom Res 2010; 68: 319-23.
7. Streiner DL, Norman GR. Health Measurement Scales. A practical guide to their development and use. New York: Oxford University Press, 2008.
8. Kitamura T. Self-rating depression scales: Some methodological issues. In: Maj M, Sartorius N (eds). Depressive Disorders, vol 1 in WPA Series Evidence and Experience in Psychiatry. New York: Wiley, 1999.
9. Terracciano A, McCrae RR, Costa PT. Factorial and construct validity of the Italian Positive and Negative Affect Schedule (PANAS). Eur J Psychol Assess 2003; 19: 131-41.
10. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. J Pers Soc Psychol 1988; 54 :1063-70.
11. Watson D, Clark LA, Carey G. Positive and negative affect and their relation to anxiety and depression disorders. J Abnorm Psychol 1988; 97: 346-53.
12. Iwata N, Roberts CR, Kawakami N. Japan-US comparison of responses to depression scale items among adult workers. Psychiatry Res1995; 58: 237-45.
13. Iwata N, Umesue M, Egashira K, Hiro H, Mizoue T, Mishima N, Nagata S. Can positive affect items be used to assess depressive disorders in the Japanese population? Psychol Med 1998; 28: 153-8.
14. Gigantesco A, Mirabella F, Bonaviri G, Morosini P. Il benessere psicologico in popolazioni del settore sanitario. Giornale Italiano di Psicopatologia 2004; 10: 315-21.
15. Spitzer R, Kroenke K, Williams J. Validation and utility of a self-report version of PRIME-MD: the PHQ Primary Care Study. JAMA 1999; 282: 1737-44.
16. Zung WW. A self-rating depression scale. Arch Gen Psychiatry 1965; 12: 63-70.
17. Demyttenaere K, De Fruyt J. Getting what you ask for: on the selectivity of depression rating scales. Psychother Psychosom 2003; 72: 61-70.
18. Clark LA, Watson D. Tripartite model of anxiety and depression: psychometric evidence and taxonomic implications. J Abnorm Psychol 1991; 100: 316-36.