Adaptation and preliminary validation of the Montgomery-Åsberg Depression Rating Scale (MADRS) using the Structured Interview Guide (SIGMA) for European Portuguese

Raquel Guiomar[1], Sara Samarra[2], Mariana Rodrigues[3], Andreia Saraiva Martins[4], Vera Martins[5], Mariana Jesus[6], Filipe Fernandes[7], Inês Alves[8], Tânia Silva[9], and Ana Ganho Ávila[10]

Abstract

The Montgomery-Åsberg Depression Rating Scale (MADRS) is considered one of the gold-standard measures to assess depression severity. To standardise the MADRS admi-

1   University of Coimbra, Center for Research in Neuropsychology and Cognitive and Behavioral Intervention, Faculty of Psychology and Educational Sciences, Coimbra, Portugal. Email: raquel.guiomar@student.fpce.uc.pt. ORCID ID: https://orcid.org/0000-0001-9611-1167

2   University of Coimbra, Faculty of Psychology and Educational Sciences, Coimbra, Portugal. Email: s.samarra2017279160@gmail.com.ORCID ID: https://orcid.org/0000-0002-5119-7099

3   University of Coimbra, Faculty of Psychology and Educational Sciences, Coimbra, Portugal. Email: uc2017270131@student.uc.pt. ORCID ID: https://orcid.org/0000-0001-5192-1668

4   University of Coimbra, Faculty of Psychology and Educational Sciences, Coimbra, Portugal. Email: afsm21@hotmail.com. ORCID ID: https://orcid.org/0000-0003-1633-8658

5   Department of Psychiatry, Coimbra Hospital and University Centre, Coimbra, Portugal. Email: vera.martins@chuc.min-saúde.pt. ORCID ID: https://orcid.org/0000-0001-7669-5724

6   Department of Psychiatry, Coimbra Hospital and University Centre, Coimbra, Portugal. Faculty of Medicine of the University of Coimbra, Coimbra, Portugal Email: uc43888@uc.pt. ORCID ID: https://orcid.org/0000-0002-3457-4454

7   Instituto Superior Manuel Teixeira Gomes, Portimão, Portugal. HPA Health Group, Portimão, Portugal. Email: filipefer-nandes@grupohpa.com. ORCID ID: https://orcid.org/0000-0001-5454-0174

8   University of Algarve, Faculty of Humanities and Social Sciences, Faro, Portugal. Email: a51761@ualg.pt. ORCID ID: https://orcid.org/0000-0002-9227-3481

9   Department of Psychiatry, Coimbra Hospital and University Centre, Coimbra, Portugal. Email: 10627@chuc.min-saude.pt. ORCID ID: https://orcid.org/0000-0002-6467-3875

10   University of Coimbra, Center for Research in Neuropsychology and Cognitive and Behavioral Intervention, Faculty of Psychology and Educational Sciences, Coimbra, Portugal. Email: ganhoavila@fpce.uc.pt. ORCID ID: https://orcid.org/0000-0001-5355-6166

nistration, a structured interview was developed (SIGMA). This study aims to translate and validate the SIGMA for European Portuguese. Twenty patients (80% women) were interviewed by ten dyads of raters (trained clinical psychologists and psychiatrists) using the European Portuguese version of the MADRS and its structured interview SIGMA. There was no significant difference in the total MADRS score between raters (interviewers and observers). The intraclass correlation for the total score between raters using the SIGMA was excellent ($r = .98$; $p < .001$). All items had excellent to good item-level intraclass correlation, and the internal consistency by rater role was good. The European Portuguese version of the SIGMA showed good preliminary psychometric properties (reliability and internal consistency). Our results suggest that the SIGMA is a useful and robust interview guide for assessing the ten depression symptoms in the MADRS, regardless of the rater clinical background.

Keywords: MADRS; SIGMA; depression; structured interview; validation studies

Adaptação e validação preliminar da entrevista estruturada (SIGMA) da Escala de Avaliação de Depressão de Montgomery-Åsberg (MADRS) para o Português Europeu

Resumo

A Escala de Avaliação de Depressão de Montgomery-Åsberg (MADRS) é considerada uma das medidas de excelência para avaliar a severidade da depressão. De forma a padronizar a administração da MADRS, foi desenvolvida uma entrevista estruturada (SIGMA). Assim, este estudo tem como objetivo traduzir e validar a SIGMA para o português europeu. Vinte participantes (80% mulheres) foram entrevistados por 10 díades de avaliadores experientes (psicólogos clínicos e psiquiatras), utilizando a versão em português europeu da MADRS e a entrevista estruturada SIGMA. Não foram encontradas diferenças significativas na pontuação total da MADRS entre os pares de avaliadores (entrevistadores e observadores). A correlação intraclasse entre avaliadores para a pontuação total da MADRS com a SIGMA foi excelente ($r = 98$; $p < .001$). Os valores de correlação intraclass ao nível do item variaram entre excelentes a bons, e a consistência interna por tipo de avaliador (entrevistador ou observador) foi também boa. A versão em português europeu da SIGMA apresentou boas propriedades psicométricas preliminares (fiabilidade e consistência interna). Os nossos resultados sugerem que a SIGMA é um guia de entrevista útil e robusto para avaliar os 10 sintomas de depressão da MADRS, independentemente da formação clínica do avaliado.

Palavras-chave: MADRS; SIGMA; depressão; entrevista estruturada; estudos de validação

INTRODUCTION

Depression is a heterogeneous diagnosis that requires robust and validated instruments for its standardised assessment. The Montgomery-Åsberg Depression Rating Scale (MADRS) is a rating scale for depression, created in 1979, that aimed to improve sensitivity and accuracy concerning clinical change estimates in patients with depression (McDowell, 2006). The MADRS has been used to assess depression severity in several studies targeting, for example, major depression disorder (Chen et al., 2020; Gaewlai et al., 2018; Mennen et al., 2020; Shi et al., 2020), bipolar disorder (Bauer et al., 2016; Damborská et al., 2019; Kamarsu et al., 2020), psychosis (Edwards et al., 2019; Montag et al., 2020), obsessive-compulsive disorder (Mosley et al., 2021), depression after stroke (Goodin et al., 2019; Verberne et al., 2019), addiction (Dandaba et al., 2020), and borderline personality disorder (Chanen et al., 2015). Many studies in distinct contexts indicate that the MADRS is a valid, reliable, and sensitive instrument to change (Bondolfi et al., 2010; Duarte-Guerra et al., 2016; Fernandes et al., 2019). In addition, the MADRS has been translated, culturally adapted, and validated to many languages, namely Japanese (Takahashi et al., 2004), German (Schmidtke et al., 1988), Thai (Satthapisit et al., 2007), Spanish (Lobo et al., 2002), Swedish (Ntini et al., 2020), French (Peyre et al., 1989), Malay (Yee et al., 2015), and Brazilian Portuguese (Fernandes et al., 2019).

In their studies, Montgomery and Åsberg (Åsberg et al., 1978) noticed differences in the clinical effectiveness between antidepressant drugs, however, the available assessment instruments could not accurately measure the clinical changes that occurred over time across various psychiatric symptoms. Therefore, aiming to distinguish the treatment effects of several antidepressants in clinical trials, the authors developed the Comprehensive Psychopathological Rating Scale (CPRS; Åsberg et al., 1978).

To select the items and build the MADRS, the authors used the CPRS to collect data about the clinical status of the patients diagnosed with depression that participated in several clinical trials of antidepressant drugs. In the inclusion process, the authors considered only patients with a primary diagnosis of depression (Feighner et al., 1972), although variability regarding the characteristics of patients was desired. The sample consisted of endogenous and reactive, psychotic and non-psychotic, bipolar and unipolar out-patients and in-patients, ranging from 18 to 69 years old. Ultimately, the authors selected ten items, which aligned with the fundamental symptoms of depression.

Several studies have compared the MADRS to other similar instruments. For instance, Heo et al. (2007) showed significant and high correlations between MADRS and the Hamilton Rating Scale for Depression (HAM-D) in a sample of

elderly patients. Similarly, Fernandes et al. (2019) showed moderate to high correlations between the MADRS and the HAM-D in a sample of adult out-patients diagnosed with major depressive disorder. Svanborg and Åsberg (2001) showed high correlations between the MADRS and the Beck Depression Inventory (BDI) in a sample of 37 in-patients diagnosed with anxiety and mood disorders. In Kroenke et al (2001) and Englbrecht et al. (2017) studies, the MADRS was also highly correlated both with the BDI-II and the Patient Health Questionnaire (PHQ-9) in a sample of patients with rheumatoid arthritis. Bjelland et al. (2002) showed a high and significant correlation between the MADRS and the Bipolar Depression Rating Scale (BDRS) in a sample of patients diagnosed with bipolar disorder. Khosravani et al. (2022) and Kjaergaard et al. (2014) showed that the MADRS was moderately correlated with the BDI-II and HADS-D. Overall, these results show MADRS's moderate to high correlations with other instruments assessing depression, across samples and diagnoses, indicating high convergent validity.

In comparison to the HAM-D, the MADRS time of administration is shorter, which is an advantage to consider in clinical trials (Iannuzzo et al., 2006). Concerning the MADRS internal structure, it differs from HAM-D in several aspects. In HAM-D, although the authors aimed for a multidimensional scale, previous studies have shown inconsistent results, varying between two to eight factors (Addington et al., 1996; Akdemir et al., 2001; O'Brien & Glaudin, 1988; Steinmeyer & Möller, 1992). Similarly, the MADRS factorial studies were not consistent with the number of factors varying from one to four (Galinowski & Lehert, 1995; Rocca et al., 2002; Serretti et al., 1999). Nevertheless, the MADRS shows increased consistency regarding the factorial structure across studies, when compared with the HAM-D (Iannuzzo et al., 2006).

Moreover, the MADRS was designed to provide a sensitive and accurate estimate of change over time (Åsberg et al., 1978). Indeed, the study by Mulder et al. (2003) showed that the MADRS was more sensitive to changes in depressive symptoms than the HAM-D across six weeks of treatment. Another important feature of the MADRS is that it can be applied by psychiatrists and non-psychiatrists and has adequate psychometric properties. Namely, Davidson et al. (1986) reported values between .57 and .76 for inter-observer correlation and .76 for the total score. Furthermore, internal consistency was significant for seven of the ten items correlated with the score of remaining items [except for the items "Reduced Sleep" (.26), "Reduced Appetite" (.12), and "Suicidal Thoughts" (.29)].

When the MADRS was first published, there were no recommended questions for clinicians to use when collecting and rating the required information per item. However, adopting a structured interview guide is known to increase reliability on rating scales that are alike (Williams & Kobak, 2008), facilitating that the same

information is gathered across patients and clinicians, consequently increasing inter-rater reliability (Williams, 1988). Furthermore, a structured interview provides novel raters with unambiguous instructions, aiding their training by providing them with a range of questions derived from expert interviewers (Williams & Kobak, 2008). To address this and to increase inter-rater reliability, Williams and Kobak (2008) developed the structured interview guide for the MADRS (SIGMA; Williams & Kobak, 2008).

Although, currently, the MADRS is one of the gold-standard measures used in research in mental health across the globe (Hengartner et al., 2020), its structured interview still needed to be translated and validated to European Portuguese. This study aims to fill this gap and complement the available MADRS with a translated and validated version of its structured interview for European Portuguese-speaking professionals.

METHOD

Participants

Twenty patients (16 women) between 38 and 73 years old (mean ± SD: 54.26 ± 8.94 years) were recruited from two distinctive healthcare units (from the centre and south of Portugal) to be interviewed. Across patients, depression severity was heterogeneous by the time of the interview, with total scores in the MADRS ranging between 3 and 40 (mean ± SD: 22.25 ± 9.07) and half of participants scoring over 23. The socio-demographic and clinical features of the sample are depicted in Table 1.

Table 1
Frequencies for the demographic characteristics of the patients (N = 20)

Gender	
Male	4
Female	16
Age Range	
20-39	1
40-49	5
50-59	9
>60	5

Table 1 (Continued)

Frequencies for the demographic characteristics of the patients (N = 20)

Education	
12th grade or below	17
Bachelor's Degree	3
Diagnosis (Health unit records)	
Depression and anxiety	16
Depression and pain	2
Bipolar Disorder (Type II)	1
Other	1

In one of the health units, the two raters were both psychologists with distinctive professional experience and worked together throughout the study. In the other health unit raters were senior psychiatrists (*n* = 3) and residents of psychiatry (*n* = 3) and the pairs were organized according to the clinical schedule of the team. Description of the raters' professional and socio-demographic profiles are depicted in Table 2.

Table 2

Frequencies for the demographic characteristics of the raters (N = 8)

Gender	
Male	2
Female	6
Age range	
20-39	7
40-49	1
Education	
Bachelor's Degree	4
Master's Degree	4
Profession	
Medicine	6
Psychology	2
Professional experience	
< 1 year	3
1-5 years	2
> 10 years	3
Site	
Centre	6
South	2

Procedure

Participants gave their informed consent before participating. The ethical committee of the affiliated research institution of the authors and of the health units where the participants were recruited were consulted and approved the study. We invited eight clinicians (six women) to participate in the study and to test the inter-rater reliability of the SIGMA, by conducting the MADRS interview using the SIGMA. Participants were organized in dyads, and each dyad interviewed each patient once. During the interview, one rater conducted the interview and the second observed it. Both raters scored the interview of the same patient independently. The raters were randomly assigned to the dyads and the interview roles. No prior training was offered to the interviewers and the only instruction given was to follow the guide as closely as possible, using the instructions, the hints, and the questions of the translated guide. None had previous experience with the MADRS.

The MADRS and the SIGMA translation

Three clinical psychologists with experience in depression assessment and fluent in English translated the MADRS items and the SIGMA questions to European Portuguese. We achieved the final European Portuguese version by combining the first three independent translations and through the discussion of all conflicting expressions. This version was then back-translated into English. The raters (psychologists and psychiatrists) assessed the final version, and indicated no further suggestions or changes.

Statistical analysis

We performed the analyses in SPSS (version 25.0; IBM Corp., 2017) and used descriptive and frequency statistics to describe the samples of participants and raters and estimate the intraclass correlation for the total score between raters. Values were interpreted with the scale suggested by Blacker and Endicott (2000) (> .80 = excellent, between .70 and .80 = good, between .50 and .70 fair, < .50 = poor). We estimated the difference in the mean MADRS total scores between raters using independent samples t-test. For internal consistency reliability estimations, we used Cronbach's alpha. Finally, we also examined the interrater reliability at the item-level, and the inter-item and item-total correlations. Correlation coefficients were interpreted according to the criteria suggested by Pestana and Gageiro (2005). That is, correlation coefficients less than .20 reflect a very low association between variables; between .21 and .39 points correspond to a low association; between .40

and .69 to moderate; between .70 and .89 are elevated; and higher than .90 are very high. For all analyses, we assumed a minimum significance level of .05.

RESULTS

We found no significant differences between the mean MADRS scores obtained by the raters who conducted the interview (M = 21.8, SD = 9.31) and the raters who observed the interview (M = 22.7, SD = 9.03, $t(38)$ = -0.31, p = .758). The intraclass correlation (ICC) between raters for the total score was excellent (r = .98, p < .001, 95% CI [.94, .99]). The internal consistency by rater role was also good, both for raters who were interviewers (α = .80) and for raters who were observers (α = .74). At the item level, the ICC estimates were good to excellent with all items showing excellent interrater reliability (r > .80, p < .01) except item 4 ("Reduced Sleep") that showed a lower ICC value (r = .70, p < .01). Regarding inter-item and inter-total correlations, item 5 and the total score were very high, item 4 was moderate, and the remaining items presented high correlations. Detailed Intraclass Correlations, Cronbach's alphas, Inter-Item and Item-Total Correlations can be found in Table 3.

Table 3
Intraclass correlations, Cronbach's alphas, inter-item and item-total correlations between raters on individual items of the MADRS

	Intraclass Correlation	Cronbach's alpha	Inter-Item Correlation	Item-Total Correlation
1. Reported Sadness	.89***	.89	.81	.81
2. Apparent Sadness	.92***	.92	.85	.85
3. Inner Tension	.89***	.89	.80	.80
4. Reduced Sleep	.70**	.70	.54	.54
5. Reduced Appetite	.95***	.95	.91	.91
6. Concentration Difficulties	.83***	.82	.71	.71
7. Lassitude	.89***	.91	.83	.83
8. Inability to Feel	.92***	.92	.86	.86
9. Pessimistic Thoughts	.92***	.92	.85	.85
10. Suicidal Thoughts	.92***	.91	.84	.84
Total	.98***	.98	.96	.96

*Note. *p < .05; **p < .01; ***p < .001; Independent Interviews, 10 dyads of raters; Sample size: N = 20*

DISCUSSION

To the best of our knowledge, this is the first translation, adaptation, and validation of the MADRS' structured interview guide for European Portuguese. Ten dyads assessed the MADRS independently (one interviewer and one observer), and a total of 20 patients from the south and centre of Portugal were included in the study. Although our sample was small, internal consistency showed positive preliminary results.

The internal consistency in our study was within the range reported in earlier studies of adult populations, despite the differences in study populations, ages, methods, and settings (Fernandes et al., 2019; Lobo et al., 2002; Ntini et al., 2020; Takahashi et al., 2004). In contrast with other studies where assessments depended on the rater background (psychiatrists, psychologists, students, and psychiatric nurses; Schmidtke et al., 1988), the values we found for inter-rater reliability were good regardless of the experience level or clinical background of the rater (psychology or medicine). This indicates the robustness, uniqueness and usefulness of the measure. These results suggest that the SIGMA might contribute to standardise the way practitioners collect information regarding depression severity. The expected increased inter-rater reliability in the MADRS with the addition of the SIGMA comes in line with what was found when a structured interview was added to the administration of Hamilton Depression and Anxiety Scales (Bruss et al., 1994; Williams, 1988).

Despite the overall high inter-item and inter-total correlations, "Reduced sleep" showed just moderate correlation coefficients. This result is in line with a previous MADRS validation where "Reduced sleep", "Reduced Appetite" and "Suicidal thoughts" correlated poorly with the total score (Davidson et al., 1986). This may be explained by a simplified vision of depression, a complex and diverse clinical condition, the frequently presents with other comorbidities. Therefore, looking at the diagnosis of depression as a discrete phenomenon according to defined and restrictive inclusion or exclusion criteria, that have to be present simultaneously with a pre-defined time-duration, overlooks phenomenological experience of symptoms and its relationships (for an in-depth discussion, see Research Domain Criteria - RDoC).

Our results should be interpreted considering some limitations. We did not include another measure of depression, so it was impossible to test the convergent validity. Concerning the sampling method, we used a convenience sample, which might reduce the generalizability of the results. On the other hand, we collected data in two clinical settings, which might contribute to the greater representability of the collected data. Another limitation of the current study was the absence of test-retest

reliability analysis. Future studies should recruit larger and more representative samples and test the convergent validity and the temporal stability of the results.

CONCLUSIONS

This preliminary study provides evidence for the validity and reliability of the European Portuguese version of the SIGMA for assessing depression severity in Portuguese patients. The study's findings suggest that the SIGMA is a useful and robust interview guide for assessing the ten depression symptoms present in the MADRS, regardless of the rater clinical background and professional experience. Additionally, it contributes with a useful tool to guide the administration of the MADRS for European Portuguese speaking clinicians and researchers, offering a better standardized assessment. The study results could additioanlly inform the development of standardized assessments for depression in other languages and cultures, allowing for cross-cultural studies and helping to ensure consistency and comparability across different settings.

ACKNOWLEDGMENTS

REFERENCES

Addington, D., Addington, J., & Atkinson, M. (1996). A psychometric comparison of the Calgary Depression Scale for Schizophrenia and the Hamilton Depression Rating Scale. *Schizophrenia Research, 19*(2-3), 205-212. https://doi.org/10.1016/0920-9964(95)00070-4

Akdemir, A., Türkçapar, M. H., Orsel, S. D., Demirergi, N., Dag, I., & Ozbay, M. H. (2001). Reliability and validity of the Turkish version of the Hamilton Depression Rating Scale. *Comprehensive Psychiatry, 42*(2), 161-165. https://doi.org/10.1053/comp.2001.19756

Åsberg, M., Montgomery, S. A., Perris, C., Schalling, D., & Sedvall, G. (1978). A comprehensive psychopathological rating scale. *Acta Psychiatrica Scandinavica. S271*, 5-27. https://doi.org/10.1111/j.1600-0447.1978.tb02357.x

Bauer, I. E., Wu, M. J., Meyer, T. D., Mwangi, B., Ouyang, A., Spiker, D., Zunta-Soares, G. B., Huang, H., & Soares, J. C. (2016). The role of white matter in personality traits and affective processing in bipolar disorder. *Journal of Psychiatric Research, 80*, 64-72. https://doi.org/10.1016/j.jpsychires.2016.06.003

Bjelland, I., Dahl, A. A., Haug, T. T., & Neckelmann, D. (2002). The validity of the Hospital Anxiety and Depression Scale. An updated literature review. *Journal of Psychosomatic Research, 52*(2), 69-77. https://doi.org/10.1016/s0022-3999(01)00296-3

Blacker, D., & Endicott, J. (2002). Psychometric properties: concepts of reliability and validity. In A. J. Rush, H. A. Pincus, M. B. First, D. A. Zarin, D. Blacker & J. Endicott. *Handbook of psychiatric measures.* American Psychiatric Association, (pp. 7-14).

Bondolfi, G., Jermann, F., Rouget, B. W., Gex-Fabry, M., McQuillan, A., Dupont-Willemin, A., Aubry, J.-M., & Nguyen, C. (2010). Self- and clinician-rated Montgomery-Åsberg Depression Rating Scale: evaluation in clinical practice. *Journal of Affective Disorders, 121*(3), 268-272. https://doi.org/10.1016/j.jad.2009.06.037

Bruss, G. S., Gruenberg, A. M., Goldstein, R. D., & Barber, J. P. (1994). Hamilton Anxiety Rating Scale Interview guide: Joint interview and test-retest methods for interrater reliability. *Psychiatry Research, 53*(2), 191-202. https://doi.org/10.1016/0165-1781(94)90110-4

Chanen, A., Jackson, H., Cotton, S. M., Gleeson, J., Davey, C. G., Betts, J., Reid, S., Thompson, K., & McCutcheon, L. (2015). Comparing three forms of early intervention for youth with border-line personality disorder (the MOBY study): Study protocol for a randomised controlled trial. *Trials, 16*(1), 1-10. https://doi.org/10.1186/s13063-015-1001-x

Chen, J. C. C., Sumner, R. L., Krishnamurthy Naga, V., Hoeh, N., Ayeni, H. A., Singh, V., Sundram, F., Campbell, D., & Muthukumaraswamy, S. (2020). A randomised, double-blind, active placebo--controlled, parallel groups, dose-response study of scopolamine hydrobromide (4-6 μg/kg) in patients with major depressive disorder. *Trials, 21*(1), 157. https://doi.org/10.1186/s13063-020-4089-6

Damborská, A., Piguet, C., Aubry, J. M., Dayer, A. G., Michel, C. M., & Berchio, C. (2019). Altered Electroencephalographic resting-state large-scale brain network dynamics in euthymic bipolar disorder patients. *Frontiers in Psychiatry, 10*, 826. https://doi.org/10.3389/fpsyt.2019.00826

Dandaba, M., Serra, W., Harika-Germaneau, G., Silvain, C., Langbour, N., Solinas, M., Noël, X., Jaafari, N., & Chatard, A. (2020). Predicting relapse in patients with severe alcohol use disorder: The role of alcohol insight and implicit alcohol associations. *Addictive Behaviors, 107*(April), 106433. https://doi.org/10.1016/j.addbeh.2020.106433

Davidson, J., Turnbull, C. D., Strickland, R., Miller, R., & Graves, K. (1986). The Montgomery-Åsberg Depression Scale: reliability and validity. *Acta Psychiatrica Scandinavica, 73*(5), 544-548. https://doi.org/10.1111/j.1600-0447.1986.tb02723.x

Duarte-Guerra, L. S., Gorenstein, C., Paiva-Medeiros, P. F., Santo, M. A., Lotufo Neto, F., & Wang, Y.-P. (2016). Clinical utility of the Montgomery-Åsberg Depression Rating Scale for the detection of depression among bariatric surgery candidates. *BMC Psychiatry, 16*, 119. https://doi.org/10.1186/s12888-016-0823-8

Edwards, C. J., Garety, P., & Hardy, A. (2019). The relationship between depressive symptoms and negative symptoms in people with non-affective psychosis: A meta-analysis. *Psychological Medicine, 49*(15), 2486-2498. https://doi.org/10.1017/S0033291719002381

Englbrecht, M., Alten, R., Aringer, M., Baerwald, C. G., Burkhardt, H., Eby, N., Fliedner, G., Gauger, B., Henkemeier, U., Hofmann, M. W., Kleinert, S., Kneitz, C., Krueger, K., Pohl, C., Roske, A. E., Schett, G., Schmalzing, M., Tausche, A. K., Peter Tony, H., & Wendler, J. (2017). Validation of standardized questionnaires evaluating symptoms of depression in rheumatoid arthritis patients: Approaches to screening for a frequent yet underrated challenge. *Arthritis Care & Research, 69*(1), 58-66. https://doi.org/10.1002/acr.23002

Feighner, J. P., Robins, E., Guze, S. B., Woodruff, R. A. J., Winokur, G., & Munoz, R. (1972). Diagnostic criteria for use in psychiatric research. *Archives of General Psychiatry, 26*(1), 57-63. https://doi.org/10.1001/archpsyc.1972.01750190059011

Fernandes, F., Carneiro, A., Campos, R., Soeiro-De-souza, M., Barros, V., & Moreno, R. (2019). SIGMA-VB: Validity and reliability of the Brazilian Portuguese version of the Montgomery-Åsberg depression rating scale using the structured interview guide for the MADRS. *Brazilian Journal of Psychiatry, 41*(4), 297-302. https://doi.org/10.1590/1516-4446-2018-0105

Gaewlai, P., Vallibhakara, S. A., Perkins, K. M., & Chantra, M. (2018). Original Article Comparative Effectiveness of Active Group Music Therapy in Major Depressive Disorder Compared to Standard Group Psychotherapy : A Randomized Controlled Trial. *Journal of the Medical Association of Thailand, 101*(4), 529-535.

Galinowski, A., & Lehert, P. (1995). Structural validity of MADRS during antidepressant treatment. *International Clinical Psychopharmacology, 10*(3), 157-161. https://doi.org/10.1097/00004850-199510030-00004

Goodin, P., Lamp, G., Vidyasagar, R., Connelly, A., Rose, S., Campbell, B. C. V., Tse, T., Ma, H., Howells, D., Hankey, G. J., Davis, S., Donnan, G., & Carey, L. M. (2019). Correlated resting-state functional MRI activity of frontostriatal, thalamic, temporal, and cerebellar brain regions differentiates stroke survivors with high compared to low depressive symptom scores. *Neural Plasticity, 2019*. https://doi.org/10.1155/2019/2357107

Hengartner, M. P., Jakobsen, J. C., Sørensen, A., & Plöderl, M. (2020). Efficacy of new-generation antidepressants assessed with the Montgomery-Åsberg Depression Rating Scale, the gold standard clinician rating scale: A meta-analysis of randomised placebo-controlled trials. *PloS One, 15*(2), e0229381. https://doi.org/10.1371/journal.pone.0229381

Heo, M., Murphy, C. F., & Meyers, B. S. (2007). Relationship between the Hamilton Depression Rating Scale and the Montgomery-Åsberg Depression Rating Scale in depressed elderly: a meta-analysis. *The American Journal of Geriatric Psychiatry, 15*(10), 899-905. https://doi.org/10.1097/JGP.0b013e318098614e

Iannuzzo, R. W., Jaeger, J., Goldberg, J. F., Kafantaris, V., & Sublette, M. E. (2006). Development and reliability of the HAM-D/MADRS interview: an integrated depression symptom rating scale. *Psychiatry Research, 145*(1), 21-37. https://doi.org/10.1016/j.psychres.2005.10.009

IBM Corp. (2017). *IBM SPSS Statistics for Windows* (25.0). IBM Corp.

Kamarsu, S., Kaufmann, C. N., Palmer, B. W., & Depp, C. A. (2020). Ecological momentary assessment of the relationships between social activity and mood in bipolar disorder. *Journal of Behavioral and Cognitive Therapy, 30*(1), 41-48. https://doi.org/10.1016/j.jbct.2020.03.018

Khosravani, V., Berk, M., Samimi Ardestani, S. M., & Sharifi Bastan, F. (2022). Confirmatory evaluation of the bipolar depression rating scale (BDRS) in a large sample of female patients with bipolar depression. *International Journal of Psychiatry in Clinical Practice, 26*(1), 85-91. https://doi.org/10.1080/13651501.2021.1904997

Kjaergaard, M., Arfwedson Wang, C. E., Waterloo, K., & Jorde, R. (2014). A study of the psychometric properties of the Beck Depression Inventory-II, the Montgomery and Åsberg Depression Rating Scale, and the Hospital Anxiety and Depression Scale in a sample from a healthy population. *Scandinavian Journal of Psychology, 55*(1), 83-89. https://doi.org/10.1111/sjop.12090

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*(9), 606-613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Lobo, A., Chamorro, L., Luque, A., Dal-Ré, R., Badia, X., Baró, E., Lacámara, C., González-Castro, G., Gurrea-Escajedo, A., Elices-Urbano, N., San-Molina, L., Ganoza-García, C., Dueñas--Herrero, R. M., Gutiérrez-Casares, J. R., Galán, F., Bustos, J., Palao-Vidal, D., Cavero-álvarez, M., Mañá-carega, S., … Martínez-Mesas, J. (2002). Validation of the Spanish versions of the Montgomery-Åsberg Depression and Hamilton Anxiety Rating Scales. *Medicina Clinica, 118*(13), 493-499. https://doi.org/10.1016/S0025-7753(02)72429-9

McDowell, I. (2006). *Measuring Health* (3rd ed.). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195165678.001.0001

Mennen, A. C., Turk-Browne, N. B., Wallace, G., Seok, D., Jaganjac, A., Stock, J., deBettencourt, M. T., Cohen, J. D., Norman, K. A., & Sheline, Y. I. (2020). Cloud-based fMRI neurofeedback to reduce the negative attentional bias in depression: a proof-of-concept study. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 6*(4), 490-497. https://doi.org/10.1016/j.bpsc.2020.10.006

Montag, C., Brandt, L., Lehmann, A., de Millas, W., Falkai, P., Gaebel, W., Hasan, A., Hellmich, M., Janssen, B., Juckel, G., Karow, A., Klosterkötter, J., Lambert, M., Maier, W., Müller, H., Pützfeld, V., Schneider, F., Stützer, H., Wobrock, T., … Gallinat, J. (2020). Cognitive and emotional empathy in individuals at clinical high risk of psychosis. *Acta Psychiatrica Scandinavica, 142*(1), 40-51. https://doi.org/10.1111/acps.13178

Mosley, P. E., Windels, F., Morris, J., Coyne, T., Marsh, R., Giorni, A., Mohan, A., Sachdev, P., O'Leary, E., Boschen, M., Sah, P., & Silburn, P. A. (2021). A randomised, double-blind, sham-controlled trial of deep brain stimulation of the bed nucleus of the stria terminalis for treatment-resistant obsessive-compulsive disorder. *Translational Psychiatry, 11*(1), 190. https://doi.org/10.1038/s41398-021-01307-9

Mulder, R. T., Joyce, P. R., & Frampton, C. (2003). Relationships among measures of treatment outcome in depressed patients. *Journal of Affective Disorders, 76*(1-3), 127-135. https://doi.org/10.1016/s0165-0327(02)00080-0

Ntini, I., Vadlin, S., Olofsdotter, S., Ramklint, M., Nilsson, K. W., Engström, I., & Sonnby, K. (2020). The Montgomery and Åsberg Depression Rating Scale – self-assessment for use in adolescents: An evaluation of psychometric and diagnostic accuracy. *Nordic Journal of Psychiatry, 74*(6), 415-422. https://doi.org/10.1080/08039488.2020.1733077

O'Brien, K. P., & Glaudin, V. (1988). Factorial structure and factor reliability of the Hamilton Rating Scale for Depression. *Acta Psychiatrica Scandinavica, 78*(2), 113-120. https://doi.org/10.1111/j.1600-0447.1988.tb06311

Pestana, M. H., & Gageiro, J. N. (2005). *Análise de dados para ciências sociais: A complementaridade do SPSS* (4th ed.). Edições Sílabo.

Peyre, F., Martinez, R., Calache, M., Verdoux, H., & Bourgeois, M. (1989). New validation of the Montgomery and Åsberg Depression Scale (MADRS) on a sample of 147 hospitalized depressed patients. *Annales Medico-Psychologiques, 147*(7), 762-767. http://www.ncbi.nlm.nih.gov/pubmed/2694887

Rocca, P., Fonzo, V., Ravizza, L., Rocca, G., Scotta, M., Zanalda, E., & Bogetto, F. (2002). A comparison of paroxetine and amisulpride in the treatment of dysthymic disorder. *Journal of Affective Disorders, 70*(3), 313-317. https://doi.org/10.1016/s0165-0327(01)00327-5

Satthapisit, S., Posayaanuwat, N., Sasaluksananont, C., Kaewpornsawan, T., & Singhakun, S. (2007). The comparison of Montgomery and Åsberg Depression Rating Scale (MADRS Thai) to Diagnostic and Statistical Manual of Mental Disorders (DSM) and to Hamilton Rating Scale for

Depression (HRSD): Validity and reliability. *Journal of the Medical Association of Thailand, 90*(3), 524-531.

Schmidtke, A., Fleckenstein, P., Moises, W., & Beckmann, H. (1988). UntersuchungenzurReliabilität und Validitäteinerdeutschen Version der Montgomery-Asberg Depression-Rating Scale (MADRS) [Studies of the reliability and validity of the German version of the Montgomery-Asberg Depression Rating Scale (MADRS)]. *Schweizer Archiv fur Neurologie und Psychiatrie, 139*(2), 51–65.

Serretti, A., Jori, M. C., Casadei, G., Ravizza, L., Smeraldi, E., & Akiskal, H. (1999). Delineating psychopathologic clusters within dysthymia: a study of 512 out-patients without major depression. *Journal of Affective Disorders, 56*(1), 17-25. https://doi.org/10.1016/s0165-0327(99)00056-7

Shi, Y., Li, J., Feng, Z., Xie, H., Duan, J., Chen, F., & Yang, H. (2020). Abnormal functional connectivity strength in first-episode, drug-naïve adult patients with major depressive disorder. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 97*, 109759. https://doi.org/10.1016/j.pnpbp.2019.109759

Steinmeyer, E. M., & Möller, H. J. (1992). Facet theoretic analysis of the Hamilton-D scale. *Journal of Affective Disorders, 25*(1), 53-61. https://doi.org/10.1016/0165-0327(92)90093-l

Svanborg, P., & Åsberg, M. (2001). A comparison between the Beck Depression Inventory (BDI) and the self-rating version of the Montgomery Åsberg Depression Rating Scale (MADRS). *Journal of Affective Disorders, 64*(2-3), 203-216. https://doi.org/10.1016/S0165-0327(00)00242-1

Takahashi, N., Tomita, K., Higuchi, T., & Inada, T. (2004). The inter-rater reliability of the Japanese version of the Montgomery-Åsberg depression rating scale (MADRS) using a structured interview guide for MADRS (SIGMA). *Human Psychopharmacology, 19*(3), 187-192. https://doi.org/10.1002/hup.576

Verberne, D., Tse, T., Matyas, T., Baum, C., Post, M., Carey, L., & van Heugten, C. (2019). Comparing Participation Outcome Over Time Across International Stroke Cohorts: Outcomes and Methods. *Archives of Physical Medicine and Rehabilitation, 100*(11), 2096-2105. https://doi.org/10.1016/j.apmr.2019.05.025

Williams, J. B. (1988). A structured interview guide for the Hamilton Depression Rating Scale. *Archives of General Psychiatry, 45*(8), 742-747. https://doi.org/10.1001/archpsyc.1988.01800320058007

Williams, J. B. W., & Kobak, K. A. (2008). Development and reliability of a structured interview guide for the Montgomery Åsberg Depression Rating Scale (SIGMA). *The British Journal of Psychiatry: The Journal of Mental Science, 192*(1), 52-58. https://doi.org/10.1192/bjp.bp.106.032532

Yee, A., Yassim, A. R. M., Loh, H. S., Ng, C. G., & Tan, K. A. (2015). Psychometric evaluation of the Malay version of the Montgomery-Åsberg Depression Rating Scale (MADRS-BM). *BMC Psychiatry, 15*(1). https://doi.org/10.1186/s12888-015-0587-6