

Statistical analysis

Initially, data were examined through simple descriptive statistics, such as means, standard deviations, frequency, and percentages. No missing data were observed. We then tested normality assumptions for each variable used in this study, including the CORE-10 items. Univariate normality was tested examining the skewness and kurtosis values (deemed as abnormal if $> |1|$ and $|3|$, respectively; Kim, 2013). All variables were normally distributed, except for age and item 6 of the CORE-10, which were strongly skewed (Table 1). Further, we tested for the presence of multivariate outliers examining (for each participant) if the chi-square value associated with the Mahalanobis distance was greater than the critical χ^2 value at a significant threshold of .05 (Tabachnick & Fidell, 2007). Twenty-nine outliers were identified and subsequently removed from the analyses. We finally examined if the CORE-10 items had a multivariate normal distribution *via* a Mahalanobis' distance-based graphical inspection (Nor, 2015), whose results were suggestive of multivariate normality.

A Confirmatory Factor Analysis (CFA) was performed on the polyconic correlation matrix to test the supposedly unidimensional structure underlying CORE-10 items (Barkham *et al.*, 2013). Parameters were estimated with the Weighted Least Squares Mean and Variance (WLSMV) estimation method for ordinal data, which is robust to non-normally distributed variables (such as in the case of Item 6). Model fit was deemed as optimal based on the following metrics and cut-offs: χ^2/DF ratio (CMIN/DF; <2), comparative fit index (CFI; ≥ 0.95), Tucker-Lewis index (TLI; ≥ 0.95), root mean square error of approximation (RMSEA; ≤ 0.08), and standardized root mean square residual (SRMR; ≤ 0.08 ; Schermelleh-Engel, Moosbrugger, & Müller, 2003; Hu & Bentler, 1999).

Internal reliability of the CORE-10 was tested through McDonald's ω and Spearman-Brown's split-half coefficient. As for construct validity, the association between CORE-10 scores and age was assessed by means of a Spearman's correlation, whilst that with sex and education through *F*-tests. Further, the association between CORE-10 and all measures used in this study (OQ45, DASS21, IIP-32, WHO-5, SWL, COMPO-12 and ECR-12) was tested via Pearson's *r* correlation coefficients, whose resulting *p*-values were corrected with the Bonferroni's method to control for the inflation of type-I error rates (Holm, 1979).

Finally, based on Aiello *et al.*'s (2022) approach, the diagnostic efficiency of the CORE-10 was examined by means of a

receiver-operating characteristics (ROC) analyses against a positive state operationalized as the combination of a DASS-21 score $>90^{\text{th}}$ percentile and a WHO-5 score $<10^{\text{th}}$ percentile of the empirical distribution (judged as indexing high levels of psychological distress and low subjective psychological well-being, respectively). Sensitivity (Se), specificity (Sp), positive and negative likelihood ratios (LR+; LR-) were derived at the optimal cut-off identified *via* Youden's *J* statistics.

All effect sizes were computed and interpreted according to guidelines (Cohen, 1988). Analyses were run on MPLUS Version 8.4 (Muthèn & Muthèn, 2017), jamovi 2.3.12 (<https://www.jamovi.org/>) and R 4.1.0 (<https://www.r-project.org/>). The significance level was set at $\alpha=0.05$.

Results

Tables 1 and 2 report descriptive statistics of all CORE-10 items and of all measures used in this study, respectively. Further, item response distributions for each item are reported in the Supplementary Materials.

The CFA showed that CORE-10 items optimally met a unidimensional structure [$\chi^2=87.792$ (35), $p<0.001$, CMIN/DF=2.508; CFI=0.983; TLI=0.978; SRMR=0.035; RMSA=0.052 (0.039-0.066)], with all items significantly loading on the underlying factor (all $p<0.004$; see Figure 1 for betas and residual variances). As for the internal consistency of this scale, McDonald's ω was 0.81, slightly benefiting only from the potential drop of item 2 (putative McDonald's $\omega=0.83$); inter-item correlations ranged from 0.33 to 0.73, without item 2 being looked at – which yielded an inter-item correlation of 0.13. Finally, Spearman-Brown's split-half coefficient was 0.79.

At $\alpha_{\text{adjusted}}=0.006$ ($\alpha_{\text{adjusted}}=0.05/k$, where *k* is equal to the number of comparisons, *i.e.* $k=8$), CORE-10 scores were positively associated with the DASS-21 [$r(548)=0.82$; $p<0.001$], OQ-45 [$r(548)=0.76$; $p<0.001$], IIP-32 [$r(548)=0.54$; $p<0.001$] and the anxiety [$r(548)=0.30$; $p<0.001$] and avoidance [$r(548)=0.27$; $p<0.001$] subscales of the ECR-12, whereas negatively with the WHO-5 [$r(548)=-0.62$; $p<0.001$], COMPO-12 [$r(548)=-0.45$; $p<0.001$] and SWL [$r(548)=-0.50$; $p<0.001$]. All effects were medium-to-large.

No associations were detected between the CORE-10 and age [$r_s(548)=-0.07$; $p=0.116$], sex [$F(1,544)=2.09$; $p=0.149$, partial $\eta^2=0.004$] or education [$F(1,538)=1.11$; $p=0.353$, partial $\eta^2=0.002$], with trivial effects.

Table 1. Descriptive for *Clinical Outcomes in Routine Evaluation-10* (CORE-10) items (N=548).

CORE-10 Items	Mean±SD	Range	Skewness	Kurtosis
Item 1	2.09±1.03	0-4	-0.142	-0.591
Item 2R*	1.65±1.10	0-4	0.223	-0.714
Item 3R*	1.43±0.85	0-4	0.327	-0.139
Item 4	1.41±1.14	0-4	0.438	-0.592
Item 5	0.76±1.03	0-4	1.148	0.264
Item 6	0.11±0.36	0-2	3.45	12.0
Item 7	1.35±1.18	0-4	0.482	-0.765
Item 8	1.10±1.11	0-4	0.745	-0.405
Item 9	1.49±1.04	0-4	0.320	-0.537
Item 10	1.34±1.16	0-4	0.463	-0.857

SD, standard deviation. *Reversed-scale item.

At an optimal cut-off of >20 ($J=0.86$), the CORE-10 showed excellent diagnostic accuracy [$AUC=0.97$; $SE=0.01$; $95\%CI(0.95, 0.99)$], as well as intrinsic ($Se=0.94$; $Sp=0.92$) and post-test diagnostics ($LR+=11.64$; $LR-=0.06$). According to such a cut-off, 13.2% of the sample was classified as obtaining an abnormal score on the CORE-10.

Discussion

This study provides Italian practitioners and clinical researchers with the adaptation and standardization of the CORE-10, a practicable screener for psychological distress that may be easily used in both clinical and research settings for routine outcome monitoring.

The Italian version of the CORE-10 proved to be i) underpinned by a mono-factorial structure, ii) internally reliable and iii) to have both convergent and divergent validity, being also iv) featured by excellent diagnostics as to the detection of psychological distress.

First, the Confirmatory Factorial Analysis evidenced a good fit to the data, suggesting a unidimensional factorial structure. That is, psychological distress as operationalized by CORE-10 may be considered a combination of three domains, namely psychological problems, functional impairments and risk to self (Barkham *et al.*, 2013). Similarly, the internal reliability of the scale was good, and this suggests that all items were closely related to each other. Of note, Item 2 (“*I have felt I have someone to turn to for support when needed*”) had the lowest loading among all others, suggesting that the variance of this item was not adequately captured by the latent dimension “distress”. In fact, there is still an open debate in the literature about whether this item should be considered a part of the psychological distress factor or, rather, a part of social well-being (*i.e.*, Fairhurst *et al.*, 2014).

Our results showed a good construct validity of the instrument. Indeed, we found a positive and significant association between the CORE-10 total score and other well-known measures of psychological distress (*i.e.* DASS-21 and OQ-45) and interpersonal problems (IIP-32), with large effects; similarly, we observed a negative and significant correlation with overall well-being (WHO-5), psychological functioning (COMPO-12) and satisfaction with life (SWL), with medium-to-large effects. Taken together, these findings suggest that an increasing level of psychological distress as evaluated by the CORE-10 corresponds

to a worse psychological functioning and quality of life. Finally, psychological distress – as indexed by greater CORE-10 total scores – was significantly and positively associated (with a small effect) with anxious and avoidant attachment, both considered a risk factor for both the onset and the maintenance of subclinical psychological suffering and psychopathology (Mikulincer & Shaver, 2012). No association was reported between CORE-10 scores and sex, age and education, suggesting that this measure is largely unrelated from sociodemographic variables, at least in our sample; thus, the possibility of a wide-spread use of the instrument in both the clinical and research setting.

Finally, the Italian version of the CORE-10 showed excellent diagnostic accuracy, as well as intrinsic and post-test diag-

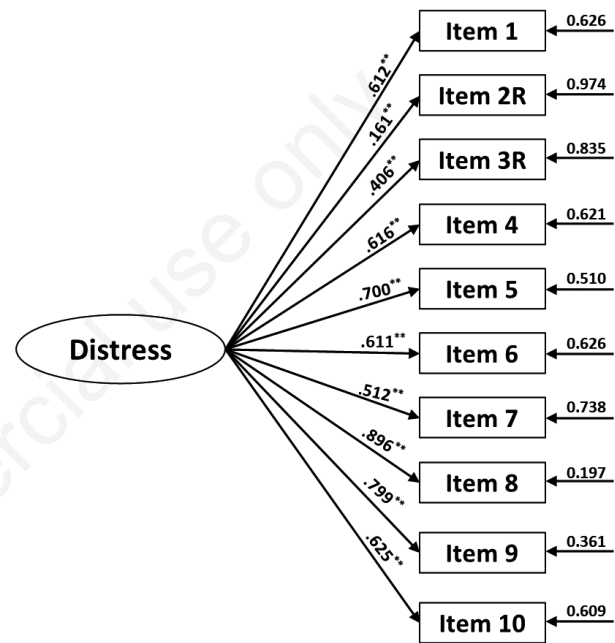


Figure 1. Standardized estimates and residual variances of the Clinical Outcomes in Routine Evaluation-Outcome Measure (N=548). Items 2 and 3 were reversed for ease of interpretation. Residual variances were computed as remainders after model estimation. ** $p<0.001$.

Table 2. Descriptive for all measures used in the study (N=548).

Variable	Mean±SD	Range
CORE-10	12.73±6.11	(1-33)
DASS-21	17.49± 10.10	(0-53)
OQ-45	55.98±19.59	(9-116)
IIP-32	38.67±15.83	(5-87)
WHO-5	13.11±4.34	(0-23)
COMPO-12	63.24±8.68	(31-84)
SWL	20.90±6.05	(5-35)
ECR-12 anxiety	4.64±1.35	(1-7)
ECR-12 avoidance	2.48±1.26	(1-7)

CORE-10, Clinical Outcomes in Routine Evaluation-Outcome Measure; DASS-21, Depression, Anxiety & Stress scale; OQ-45, Outcome Questionnaire 45; IIP-32, Inventory of Interpersonal Problems; WHO-5, World Health Organization-Five Well-Being Index; COMPO-12, Complementary Measure of Psychotherapy Outcome; SWL, Satisfaction with Life questionnaire; ECR-12, Experiences in Close Relationships scale.

nostics. An optimal cut-off of 20 was found to be critical for identifying highly distressed individuals with a poor well-being. In this regard, from a clinical perspective, it is interesting to note that more than 10% of the participants scored above the psychopathological cut-off. The prevalence of severe psychological distress in our sample was similar to that of age-matched, nationally representative samples of emerging adults in the US (11.99% in 2017; Twenge *et al.*, 2019).

Concerning the possible limitations of the present study, we may mention that we did not assess the test-retest stability or examine social desirability effects. Also, we may note that the large majority of the sample was composed by students, possibly limiting the generalizability of the results to other populations (including the clinical one).

Conclusions

To conclude, this study provides the first evidence on the psychometric robustness of the Italian version of the CORE-10 as a brief and reliable measure of psychological distress, encouraging therapists to adopt this instrument for effective routine outcomes evaluation in both the clinical and research setting.

References

- Aiello, E. N., Fiabane, E., Margheritti, S., Magnone, S., Bolognini, N., Miglioretti, M., & Giorgi, I. (2022). Psychometric properties of the Copenhagen Burnout Inventory (CBI) in Italian Physicians. *La Medicina del Lavoro*, *113*(4), e2022037. doi: 10.23749/mdl.v113i4.13219
- Barkham, M., Bewick, B., Mullin, T., Gilbody, S., Connell, J., Cahill, J., Mellor-Clark, J., Richards, D., Unsworth, G., & Evans, C. (2013). The CORE-10: A short measure of psychological distress for routine use in the psychological therapies. *Counselling & Psychotherapy Research*, *13*(1), 3-13. doi: 10.1080/14733145.2012.729069
- Barkham, M., Hardy, G. E., & Mellor-Clark, J. (Eds.). (2010). *Developing and delivering practice-based evidence: A guide for the psychological therapies*. John Wiley & Sons.
- Barkham, M., Margison, F., Leach, C., Lucock, M., Mellor-Clark, J., Evans, C., ... McGrath, G. (2001). Service profiling and outcomes benchmarking using the CORE-OM: Toward practice-based evidence in the psychological therapies. *Journal of Consulting and Clinical Psychology*, *69*(2), 184-196. doi: 10.1037/0022-006x.69.2.184
- Bottesi, G., Ghisi, M., Altoè, G., Conforti, E., Melli, G., & Sica, C. (2015). The Italian version of the Depression Anxiety Stress Scales-21: Factor structure and psychometric properties on community and clinical samples. *Comprehensive Psychiatry*, *60*, 170-181. doi: 10.1016/j.comppsy.2015.04.005
- Brugnera, A., Zarbo, C., Farina, B., Picardi, A., Greco, A., Lo Coco, G., Greco, F. (2019). Psychometric properties of the Italian version of the Experience in Close Relationship Scale 12 (ECR-12): an exploratory structural equation modeling study. *Research in Psychotherapy: Psychopathology, Process, Outcome*, *22*(3). doi: 10.4081/ripppo.2019.392
- Chiappelli, M., Coco, G. L., Gullo, S., Bensi, L., & Prestano, C. (2008). The Outcome Questionnaire 45.2. Italian validation of an instrument for the assessment of psychological treatments. *Epidemiology and Psychiatric Sciences*, *17*(2), 152-161. doi: 10.1017/s1121189x00002852
- Chui, H., Chong, E., Atzil-Slonim, D., Sahin, Z., Solomonov, N., Minges, M. V., Barber, J. P. (2021). Beyond symptom reduction: Development and validation of the Complementary Measure of Psychotherapy Outcome (COMPO). *Journal of Counseling Psychology*, *68*(5), 550-561. doi: 10.1037/cou0000536
- Cohen J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- De Jong, K., Timman, R., Hakkaart-Van Rijen, L., Vermeulen, P., Kooiman, K., Passchier, J., & Busschbach, J. V. (2014). The effect of outcome monitoring feedback to clinicians and patients in short and long-term psychotherapy: A randomized controlled trial. *Psychotherapy Research*, *24*(6), 629-639. doi: 10.1080/10503307.2013.871079
- Di Fabio, A., & Gori, A. (2016). Measuring Adolescent Life Satisfaction: Psychometric Properties of the Satisfaction with Life Scale in a Sample of Italian Adolescents and Young Adults. *Journal of Psychoeducational Assessment*, *34*(5), 501-506. doi: 10.1177/0734282915621223
- Duncan, B. L. (2012). The Partners for Change Outcome Management System (PCOMS): The Heart and Soul of Change Project. *Canadian Psychology/Psychologie canadienne*, *53*(2), 93. doi: 10.1037/a0027762
- Duncan, B. L., & Reese, R. J. (2015). The Partners for Change Outcome Management System (PCOMS) revisiting the client's frame of reference. *Psychotherapy*, *52*(4), 391-401. doi: 10.1037/pst0000026
- Elfström, M. L., Evans, C., Lundgren, J., Johansson, B., Hakeberg, M., & Carlsson, S. G. (2013). Validation of the Swedish version of the Clinical Outcomes in Routine Evaluation Outcome Measure (CORE-OM). *Clinical Psychology & Psychotherapy*, *20*(5), 447-455. doi: 10.1002/cpp.1788
- Evans, C., Connell, J., Barkham, M., Margison, F., McGrath, G., Mellor-Clark, J., & Audin, K. (2002). Towards a standardized brief outcome measure: Psychometric properties and utility of the CORE-OM. *The British Journal of Psychiatry*, *180*(1), 51-60. doi: 10.1192/bjp.180.1.51
- Evans, C., Connell, J., Barkham, M., Marshall, C., & Mellor-Clark, J. (2003). Practice-based evidence: benchmarking NHS primary care counselling services at national and local levels. *Clinical Psychology & Psychotherapy: An International Journal of Theory & Practice*, *10*(6), 374-388. doi: 10.1002/cpp.384
- Fairhurst, C., Böhnke, J. R., Gabe, R., Croudace, T. J., Tober, G., & Raistrick, D. (2014). Factor analysis of treatment outcomes from a UK specialist addiction service: Relationship between the Leeds Dependence Questionnaire, Social Satisfaction Questionnaire and 10-item Clinical Outcomes in Routine Evaluation. *Drug and Alcohol Review*, *33*(6), 643-650. doi: 10.1111/dar.12146
- Gallo, E., & Rucci, P. (2000). Supply, demand and predictive factors of psychotherapies in 10 community mental health services in Emilia Romagna. *Epidemiology and Psychiatric Sciences*, *9*(2), 103-112. doi: 10.1017/s1121189x00008290
- Harkness, J., Pennell, B. E., & Schoua-Glusberg, A. (2004). Survey questionnaire translation and assessment. *Methods for Testing and Evaluating Survey Questionnaires*, 453-473. doi: 10.1002/0471654728.ch22
- Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, *6*(2), 65-70. <http://www.jstor.org/stable/4615733>
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus

- new alternatives. *Structural Equation Modeling: a Multidisciplinary Journal*, 6(1), 1-55. doi: 10.1080/10705519909540118
- Kim, H. Y. (2013). Statistical notes for clinical researchers: Understanding standard deviations and standard errors. *Restorative Dentistry & Endodontics*, 38(4), 263. doi: 10.5395/rde.2013.38.4.263
- Lambert, M. J. (2010). *Prevention of treatment failure: The use of measuring, monitoring, and feedback in clinical practice*. Washington, DC, USA: American Psychological Association.
- Lambert, M.J., Burlingame, G.M., Umphress, V., Hansen, N.B., Vermeersch, D.A., Clouse, G.C. (1996). The reliability and validity of the Outcome Questionnaire. *Clinical Psychology and Psychotherapy*, 3(4), 249–258. doi: 10.1002/(SICI)1099-0879(199612)3:4<3C249::AID-CPP106%3E3.0.CO;2-S
- Lambert, M. J., Whipple, J. L., & Kleinstäuber, M. (2018). Collecting and delivering progress feedback: A meta-analysis of routine outcome monitoring. *Psychotherapy (Chicago, Ill.)*, 55(4), 520-537. doi: 10.1037/pst0000167
- Lambert, M. J., Whipple, J. L., Hawkins, E. J., Vermeersch, D. A., Nielsen, S. L., & Smart, D. W. (2003). Is It Time for Clinicians to Routinely Track Patient Outcome? A Meta-Analysis. *Clinical Psychology: Science and Practice*, 10(3), 288-301. doi: 10.1093/clipsy.bpg025
- Lambert, M. J., Whipple, J. L., Vermeersch, D. A., Smart, D. W., Hawkins, E. J., Nielsen, S. L., & Goates, M. (2002). Enhancing psychotherapy outcomes via providing feedback on client progress: a replication. *Clinical Psychology & Psychotherapy*, 9(2), 91-103. doi: 10.1002/cpp.324
- Lo Coco, G., Mannino, G., Salerno, L., Oieni, V., Di Fratello, C., Profita, G., & Gullo, S. (2018). The Italian Version of the Inventory of Interpersonal Problems (IIP-32): Psychometric Properties and Factor Structure in Clinical and Non-clinical Groups. *Frontiers in psychology*, 9, 341. doi: 10.3389/fpsyg.2018.00341
- Lo Coco, G., Chiappelli, M., Bensi, L., Gullo, S., Prestano, C., & Lambert, M. J. (2008). The factorial structure of the outcome questionnaire-45: a study with an Italian sample. *Clinical Psychology & Psychotherapy*, 15(6), 418-423. doi: 10.1002/cpp.601
- Lomazzi, L., Fava, E., Landra, S., D'Angelo, P., Lammoglia, M., Pazzi, E., ... & Carta, I. (1997). Psychotherapies in Lombard public mental health services: psychiatrists' and psychologists' point of view. *Epidemiology and Psychiatric Sciences*, 6(3), 184-193.
- Lomazzi, L., Fava, E., Landra, S., D'Angelo, P., Lammoglia, M., Pazzi, E., ... Carta, I. (1997). Psychotherapies in Lombard public mental health services: psychiatrists' and psychologists' point of view. *Epidemiology and Psychiatric Sciences*, 6(3), 184-193. doi: 10.1017/s1121189x00005030
- Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the depression anxiety stress scales* (2nd ed.). Sydney, N.S.W.: Psychology Foundation of Australia.
- Mikulincer, M., & Shaver, P. R. (2012). An attachment perspective on psychopathology. *World psychiatry: official journal of the World Psychiatric Association (WPA)*, 11(1), 11–15. doi: 10.1016/j.wpsyc.2012.01.003
- Miller, S. D., Duncan, B. L., Brown, J., Sparks, J. A., & Claud, D. A. (2003). The outcome rating scale: A preliminary study of the reliability, validity, and feasibility of a brief visual analog measure. *Journal of brief Therapy*, 2(2), 91-100.
- Muthén, M., & Muthén, B. O. (2017). *Mplus User's Guide*. Los Angeles, CA: Muthén & Muthén.
- Nor, A.W. (2015). The Graphical Assessment of Multivariate Normality Using SPSS. *Education in Medicine Journal*, 7(2). doi: 10.5959/eimj.v7i2.361
- Palmieri, G., Evans, C., Hansen, V., Brancaleoni, G., Ferrari, S., Porcelli, P., Reitano, F., & Rigatelli, M. (2009). Validation of the Italian version of the Clinical Outcomes in Routine Evaluation Outcome Measure (CORE-OM). *Clinical Psychology and Psychotherapy*, 16(5), 444-449. doi: 10.1002/cpp.646
- Pavot, W., & Diener, E. (1993). Review of the Satisfaction With Life Scale. *Psychological Assessment*, 5(2), 164-172. doi: 10.1037/1040-3590.5.2.164
- Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures. *Methods of Psychological Research*, 8(2), 23-74.
- Shepherd, M., Ashworth, M., Evans, C., Robinson, S. I., Rendall, M., & Ward, S. (2005). What factors are associated with improvement after brief psychological interventions in primary care? Issues arising from using routine outcome measurement to inform clinical practice. *Counselling and Psychotherapy Research*, 5(4), 273-280. doi: 10.1080/14733140600571326
- Tabachnick B. G., Fidell L. S. (2007). *Using multivariate statistics* (5th ed ed.). Boston, MA: Pearson Education, Inc.
- Topp, C. W., Østergaard, S. D., Søndergaard, S., & Bech, P. (2015). The WHO-5 Well-Being Index: a Systematic Review of the Literature. *Psychotherapy and Psychosomatics*, 84(3), 167-176. doi: 10.1159/00037658
- Twenge, J. M., Cooper, A. B., Joiner, T. E., Duffy, M. E., & Binau, S. G. (2019). Age, period, and cohort trends in mood disorder indicators and suicide-related outcomes in a nationally representative dataset, 2005–2017. *Journal of Abnormal Psychology*, 128(3), 185-199. doi: 10.1037/abn0000410
- WHO (1998). *Wellbeing Measures in Primary Health Care/The Depcare Project*. WHO Regional Office for Europe: Copenhagen.
- Zhang, Y., Hu, J., Evans, C., Jin, L. W., Wu, M. Y., Wang, C. Y., & Chen, G. P. (2020). Psychometric properties of the Chinese version of the clinical outcomes in routine evaluation-outcome measure (CORE-OM). *British Journal of Guidance and Counselling*, 48(2), 289-299. doi: 10.1080/03069885.2019.1682120