

Exploration of psychometric properties of the Italian version of the Core Young Person's Clinical Outcomes in Routine Evaluation (YP-CORE)

Rosalba Di Biase,¹ Chris Evans,² Daniela Rebecchi,³ Flavia Baccari,⁴ Anna Saltini,¹ Elena Bravi,⁵ Gaspare Palmieri,⁶ Fabrizio Starace¹

¹Department of Mental Health, Azienda Unità Sanitaria Locale, Modena, Italy; ²Department of Psychology, Cathedral Court, Vicar Lane, Sheffield, UK; ³Italian Society of Hospital and Territorial Psychology, Modena, Italy; ⁴Istituto delle Scienze Neurologiche, Bologna, Italy; ⁵Department of Mental Health Services, Trento; and Italian Society of Hospital and Territorial Psychology, Italy; ⁶Psychiatrist, Private Hospital, Villa Pineta, Modena, Italy

Correspondence: Rosalba Di Biase, Department of Mental Health, Azienda Unità Sanitaria Locale di Modena, viale Ludovico Antonio Muratori 201, 41124 Modena, Italy.
E-mail: r.dibiase@ausl.mo.it

Acknowledgements: the authors thank all participants for their contributions to the research. Thanks to SIPSOT (Italian Society of Hospital and Territorial Psychology).

Contributions: all authors contributed in study and all approved the final version.

Conflict of interests: the authors declare no potential conflict of interests; CE is a Trustee of the CORE System Trust, which holds the copyright for the YP-CORE, but derives no income from that as the instrument is free to reproduce. The study did not receive any external funding.

Availability of data and material: the authors confirm that the data supporting the findings of this study are available within the article [and/or] its supplementary materials.

Ethical approval and consent to participate: the study was designed in accordance with the tenets of the Declaration of Helsinki. All participants and their legal guardians provided written informed consent. Permission to conduct the study was granted by authorities of Azienda Unità Sanitaria Locale (AUSL) of Modena and Trento.

Consent for publication: authors consent to publish the article and also grant any third party the right to use, distribute and reproduce the article in any non-commercial medium, provided the original work is properly cited.

Borghi, L., Bonazza, F., Lamiani, G., Musetti, A., Manari, T., Filosa, M., Quattropiani, M.C., Lenzo, V., Freda, M.F., Lemmo, D., Saita, E., Cattivelli, R., Castelnuovo, G., Vegni, E., & Franceschini, C. (2021). Dreaming during lockdown: a quali-quantitative analysis of the Italian population dreams during the first COVID-19 pandemic wave. *Research in Psychotherapy: Psychopathology, Process and Outcome*, 24(2), 231-239. doi: 10.4081/ripppo.2021.554

Received for publication: 11 June 2021.
Accepted for publication: 13 June 2021.

This work is licensed under a Creative Commons Attribution Non-Commercial 4.0 License (CC BY-NC 4.0).

©Copyright: the Author(s), 2021
Licensee PAGEPress, Italy
Research in Psychotherapy: Psychopathology, Process and Outcome 2021; 24:231-239
doi:10.4081/ripppo.2021.554

ABSTRACT

There is an increasing need of outcome measures for young people suitable to evaluate treatments and routine settings. However, measures must show suitable psychometric characteristics for such use. This is the first psychometric evaluation of the Italian version of the Core Young Person Clinical Outcomes in Routine Evaluation (YP-CORE). Data are reported for a clinical sample, aged 11-17 (n=175) and non-clinical sample, aged 11-17 (n=206). Analyses included acceptability, confirmatory factor analysis, internal reliability, influence of gender and age on cut-off scores and reliable change. The YP-CORE acceptability was good, with a very high completion rate (98.7% fully completed). Internal consistency was good: the overall Cronbach's alpha value (α) equal to 0.75 (95% confidence interval=0.69-0.80). The measure was sensitive to change (Cohen $d_z=1.35$). The Italian version of the YP-CORE showed acceptable psychometric properties is suitable for use in services for young people as a change/outcome measure.

Key words: Adolescents; mental health; psychological therapy; YP-CORE questionnaire; psychometric characteristics.

Introduction

The increasing need to improve the quality of mental health services entails the use of outcome measures to evaluate treatment interventions and routine services; simultaneously, the tightened economic situation creates pressure for routine outcome exploration (Crawford, Zoha, Macdonald, Kingdon, 2017; Kilbourne *et al.*, 2018).

In services for adolescents within the Italian mental health services routine evaluation is not common but is increasing. The most popular measures currently used have been developed for other purposes. For example, several tools were designed for extremely mentally ill individuals, or were applied as screening measures or instruments to detect specific distress or specific symptoms domains (in order to provide diagnosis or to assess symptoms severity). This limits their application in the routine clinical practice (Twigg *et al.*, 2009). An outcome instrument for routine clinical use has to be quick and easily usable to detect the clinical change produced

by treatment interventions. Moreover, an outcome instrument must be a general tool useful in different clinical settings and in different mental health contexts (Barkham *et al.*, 2001; Evans *et al.*, 2002).

One of most widely used set of instruments for outcome evaluation is the *Clinical Outcomes in Routine Evaluation* (CORE) system, launched in 1998 by the CORE System Group in the United Kingdom (www.coresystemtrust.org.uk). The CORE-OM has experienced an extraordinary uptake since its creation. The CORE-OM is a 34-item self-report measure designed to evaluate the adult global level of distress. The instrument covers four domains: subjective well-being, problems/symptoms, life/social functioning and risk to self/others and there are approved translations for more than 30 languages showing satisfactory psychometric properties. The CORE system also includes the Young Person's Clinical Outcomes in Routine Evaluation (see YP-CORE; Twigg *et al.*, 2009 and link.coresystemtrust.org.uk/yp for information) derived from the CORE-OM but designed for adolescents. The aim for the YP-CORE was a brief measure, easy to complete by adolescents and easy to use in routine clinical practice. Like the CORE-OM, the YP-CORE provides a generic measure to detect a wide range of presenting problems as well overall functioning. The questionnaire provides a global mental distress index and is therefore suitable for use as an initial assessment tool and as a measure of outcomes; it cannot be used to obtain a specific disorder diagnosis. (Evans *et al.*, 2002; Twigg *et al.*, 2016; O'Reilly, Peiper, Keeffe, Illback, Clayton, 2016; Gergov *et al.*, 2017; Feixas *et al.*, 2018).

Other self-report validated measures designed for use with help-seeking young people include Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997), Health of the Nation Outcome Scales for Children and Adolescents Self Report (HoNOSCA-SR; Gowers *et al.*, 1999a; 1999b), Children Depression Inventory (CDI; Kovacs, 1985), Child Behaviour Checklist (CBCL, Achenbach, 1991). However, these each have at least one drawback compared to the YP-CORE: requiring licence payments, having a longer time frame than one week, being tuned to a relatively high level of distress/dysfunction or tending to locate items in a diagnostic/disease framework not particularly useful to the measurement of change in therapies. Most critically most of these scales are designed to identify more severe 'cases' and are not suitable for routine use in places such as school counselling services.

This study assesses psychometric characteristics of the YP-CORE Italian Version. As it is primarily intended for use to measure change in interventions for help-seeking young people the exploration is of both a routine sample of such young people attending clinical services, and, to provide the vital context, of a school attender sample.

Materials and Methods

Adaptation of the instrument

The Italian YP-CORE translation was conducted in line with current recommendations (Prakash, Shah, & Hariom, 2019) and to the specific protocol required the CORE System Trust (CST, 2015; <https://www.coresystemtrust.org.uk/translations/>) using a mixed translation procedure of forward and back translation, group review and field testing (see Yassin & Evans, 2021 for a complete account of this process). For the Italian YP-CORE independent forward translations were produced by nine mental health professionals (five female and four male). An independent back translation was then produced by a professional bilingual translator. The translations obtained were compared and the final version was reviewed by a member of CORE System Trust (CE). The quality of the translation was also verified through a group of ten Italian-speaking, teenagers aged 14-17 (four female and six male all of medium socioeconomic background). They were asked to rate each item's comprehensibility a three level scale. Eight of the adolescents rated all the items as easy to read and comprehensible (score 3); one adolescent rated item 5 as unclear (score 2) and one more rated it as not at all clear (score 1). This led to small changes to item 5 to align the language more closely to that of adolescents.

Participants in the psychometric evaluation

This was an unfunded study so the sampling frames were convenience frames based on the locations of the authors but chosen to cover the full age range of 11 to 17. Data collection took place between May 2017 and June 2018 with a clinical and a non-clinical sample. Inclusion criteria were: age 11-17 and Italian-speaking; exclusion criteria were diagnosis of psychotic disorder or intellectual disability. The clinical sample (CS) were patients consecutively recruited from two psychotherapeutic services for children and adolescents in Italy. One sample, aged 11 to 13, was from the Trento Psychological Centre for Childhood and Adolescents in the Trentino-Alto Adige province. The second sample, aged 13 to 17, was from the Modena Mental Health Service for Adolescents in the Emilia Romagna province. The combined sample included 118 female and 57 male participants (total n 175) with a wide range of diagnoses and mean age overall 14.7 years [standard deviation (SD)=1.6; female mean age 14.9, SD 1.5; male mean 14.3, SD=1.9].

The non-clinical sample (NCS) was recruited in middle (ages 11-13) and high (ages 14-17) schools in Modena. Data were collected by a psychologist in the school after explaining the study to the young people and after consent from their parents/legal guardians. All participants were fluent in Italian and there were no exclusion criteria. In total, 206 adolescents (female 129; male 77) were invited to participate and all agreed. The overall

mean age was 14.3 (SD 1.8), 14.3 for the females (SD 1.9) and 14.1 (SD 1.9) for the males.

A pre-post design was used in the clinical sample with a psychologist offering the questionnaire to adolescents at the first session and again in the last session.

Measures

The YP-CORE is a brief generic measure of psychological distress for young people, aged between 11 and 16. The questionnaire includes ten self-report items (one Risk item, one Wellbeing item, four Problems/Symptoms items, and four General Functioning items) each answered on a 5-point Likert frequency scale scored from 0 to 4 with the total score also scored 0 to 4 as the mean of all items or nine items if only one item is omitted. The full translated measure can be downloaded from <https://www.coresystemtrust.org.uk/translations/italian/>.

Statistical analysis

The approach is descriptive and exploratory recognising that the unfunded nature for the research meant that the non-clinical samples are opportunistic not rigorous random samples of the population. Bootstrapped confidence intervals (CIs) based on percentile CI extraction and 10,000 replications are reported where possible to avoid the restrictions of hypothesis testing and of distributional issues. Acceptability of YP-CORE was tested by measuring the proportion of missed items at baseline, both in clinical and non-clinical samples. Cronbach's alpha, based on baseline data, was used to evaluate internal consistency for both samples and for each gender subsample. Sensitivity to change in CS was evaluated by pre-post mean change and effect size (ES) according to Cohen's d_1 and d_z (Cohen, 1988). Following the Jacobson and Truax method, Cronbach's alpha value of the combined sample (clinical and non-clinical) were used to calculate the reliable change index (RCI) for the clinical sample; then the RCI approach was used to categorize change for individuals as more or

less than would be expected by unreliability of measurement alone (Jacobson and Truax, 1991). Cut-off values for clinically significant change (CSC) were calculated using Jacobson and Truax's method 'c' (Jacobson & Truax, 1991), considering means and standard deviations from CS and NCS. Confirmatory factor analysis (CFA) of the combined baseline data tested a one factor model, a two factor model found in Twigg *et al.* (2009) with factors for the seven negatively cued items and the three positively cued items and this was contrasted with the single factor model using diagonally weighted least squares (DWLS), model comparison used the Satorra-Bentler scaled chi squared method (Satorra & Bentler, 2001) using Lavaan (2012) v. 0.6.7 (Rosseel, 2012). Analyses were carried out using SAS Version 9.1 (SAS Institute, 2011) and R version 4.0.2 (R CORE Team, 2020).

Results

Acceptability

Over the combined samples 376 participants (98.7%) completed the entire questionnaire. In the clinical sample, all the 175 baseline YP-CORE forms were fully completed, whereas in the NCS 201 participants (97.6%) completed the entire questionnaire. The most commonly omitted items were item 3 (*I've felt able to cope when things go wrong*) and item 7 (*My problems have felt too much for me*).

Mean scores and effect of gender and age

Means and standard deviations of YP-CORE scores by age group and gender at baseline for the two samples are shown in Table 1 and the distributions of scores by gender and sample are shown in Figure 1. Figure 2 shows the relationship between scores and age by gender and sample (in Figure 2, points are 'jittered' horizontally *i.e.* small random adjustments to age are made to prevent overprinting).

Table 1. Means and standard deviations by gender and age group for YP-CORE pre-intervention in clinical and non-clinical samples.

Groups	No.	Min	Clinical sample				N	Min	Non-clinical sample			
			Mean	SD	CI (mean)	Max			Mean	SD	CI (mean)	Max
Female	118	0.4	1.99	0.70	[1.86:2.11]	3.3	127	0.0	1.04	0.46	[0.96:1.13]	2.1
Male	57	0.5	1.68	0.59	[1.54:1.82]	3.2	77	0.1	0.87	0.42	[0.78:0.97]	2.0
11-14	72	0.5	1.72	0.67	[1.58:1.87]	3.3	107	0.0	0.96	0.47	[0.87:1.04]	2.1
15-17	103	0.4	2.00	0.67	[1.86:2.13]	3.3	97	0.1	1.00	0.44	[0.91:1.10]	2.1
11-14 F	44	0.6	1.82	0.72	[1.62:2.03]	3.3	63	0.0	1.01	0.49	[0.89:1.14]	2.1
11-14 M	28	0.5	1.57	0.55	[1.36:1.79]	3.2	44	0.1	0.88	0.42	[0.75:1.01]	2.0
15-17 F	74	0.4	2.09	0.67	[1.93:2.25]	3.3	64	0.4	1.07	0.44	[0.97:1.18]	2.1
15-17 M	29	0.6	1.78	0.61	[1.55:1.99]	2.9	33	0.1	0.86	0.42	[0.72:1.00]	1.8

SD, standard deviation; CI, confidence interval.

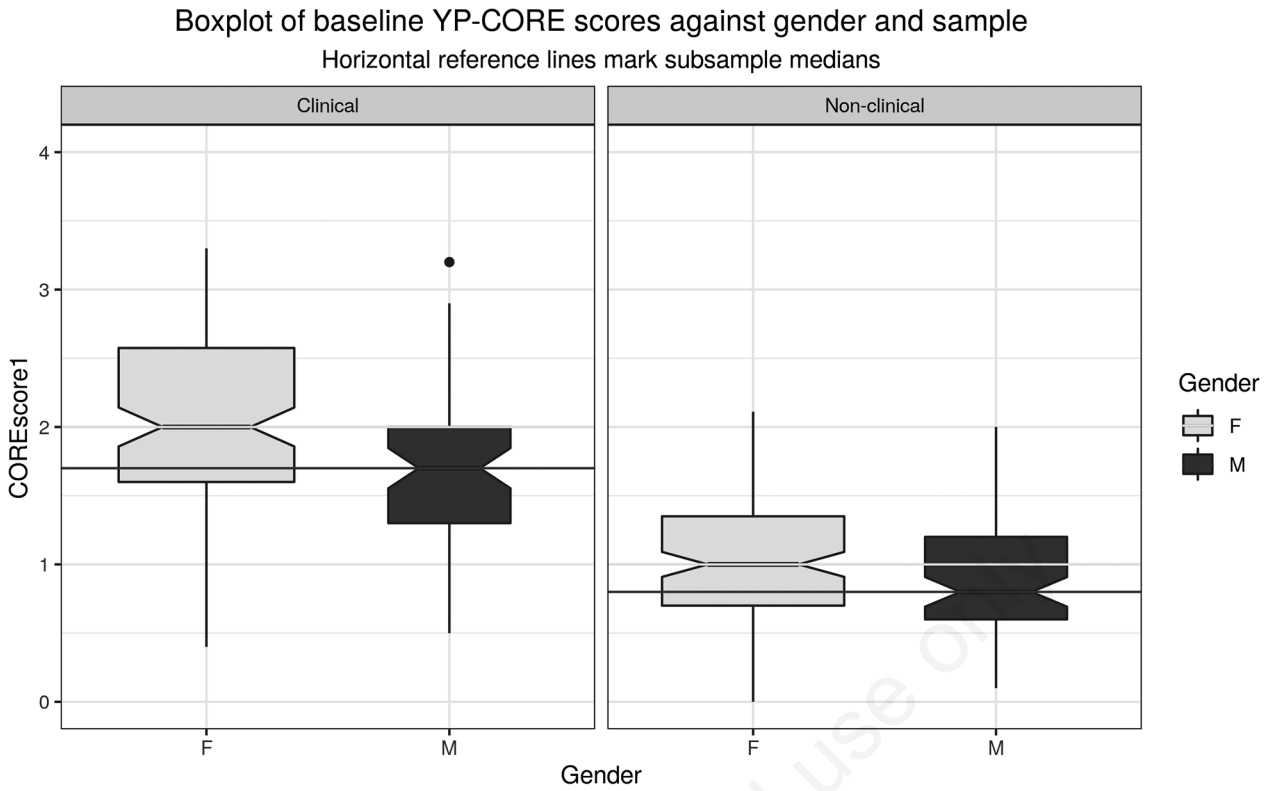


Figure 1. Relationship between YP-CORE scores and age by sample.

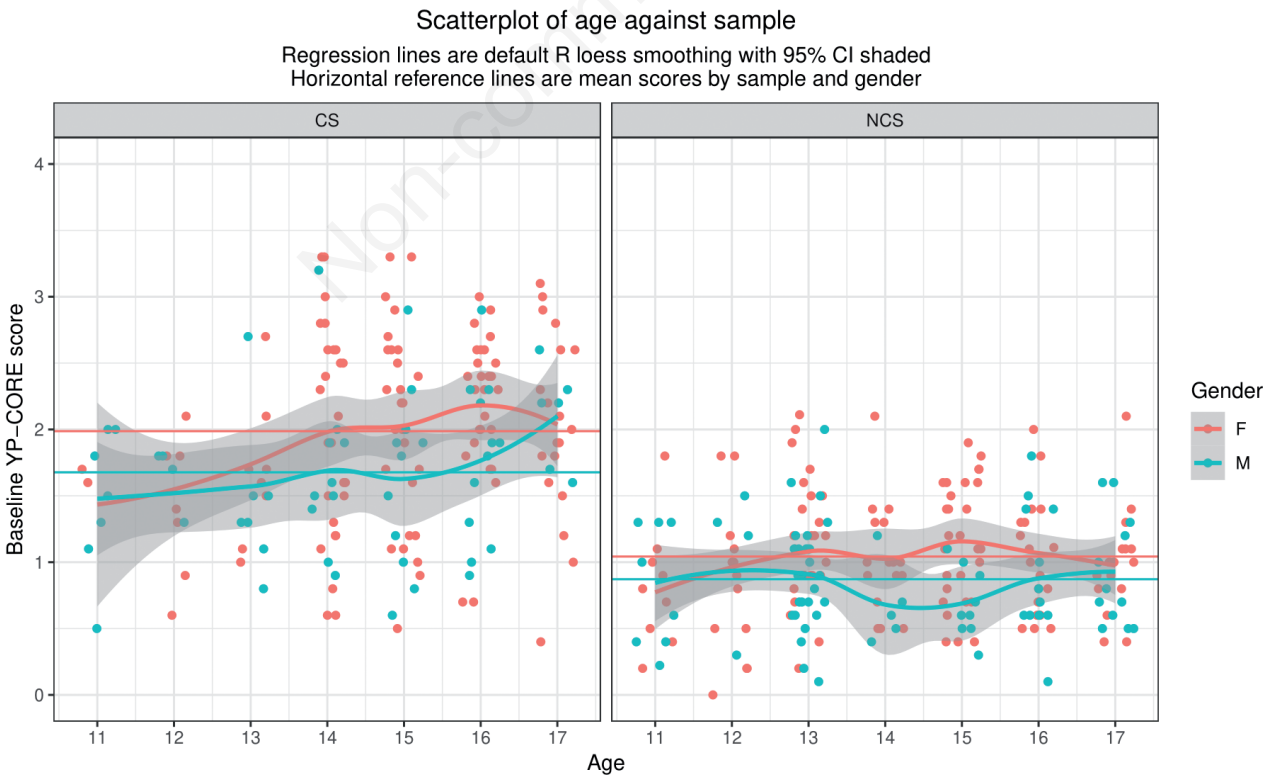


Figure 2. Relationship between YP-CORE score and age. CI, confidence interval; CS, clinical sample; NCS, non-clinical sample.

Internal consistency

Cronbach's alpha values are shown in Table 2. There are clear differences by gender and age but the values are acceptable though the value of 0.69 for the males in the 11 to 14 age group is marginal.

Sensitivity to change

Post-intervention data were available for 74 participants of the CS (42.3%). No difference was found in age and gender between patients in CS post-treatment sample and the whole CS sample. YP-CORE scores were lower in the follow-up assessment than at baseline as shown in Table 3. The mean pre-post difference for the whole sample was 0.74 (95% CI=0.60-0.85; SD=5.5; $P<0.001$), suggesting good sensitivity to change. Effect size as Cohen's d was 1.11 in the whole sample, 1.10 in the females and 1.22 in the males (Figures 3-5).

Table 4 and Figure 5 show the clinically significant change (CSC) cutting points based on these data with 95% CIs. It can be seen clearly in Figure 3 that although the confidence intervals are wide for the small subgroupings, nevertheless it is clear that different CSCs are needed by gender and age.

Factor structure

Inter-item correlations across the $n=376$ full baseline sample with complete data ranged from 0.04 to 0.63, and 41 out of 45 correlations were statistically significant. The one-factor solution showed a good fit for the YP-CORE, raw $\chi^2(35)=33.9$, $P=0.52$; CFI=1.00; TLI=1.00, RMSEA<0.001, 90% CI for RMSEA=0.00-0.036. Standardized factor loadings are reported in Table 5. Despite

the excellent fit, adding a second factor with the negatively cued items loading on the first factor and the negatively cued items loading on the second factor showed an even better fit: raw $\chi^2(34)=22.8$, $P=0.93$; CFI=1.00; TLI=1.01, RMSEA<0.001, 90% CI for RMSEA=0.00-0.012 and the difference was statistically significant robust raw $\chi^2(1)=14.2$, $P=0.0002$. Both one and two factor solutions are shown in Table 6 and Figure 6.

Discussion

This study reports the first evaluation of psychometric properties of the Italian version of YP-CORE. The measure showed excellent acceptability with no missed items in the clinical sample and very few in the non-clinical sample with only two of the 376 responses (0.5%) not proratable, consistent with findings in the UK, Finland and Spain (Twigg *et al.*, 2016; Gergov *et al.*, 2017; Feixas *et al.*, 2018). This suggests that YP-CORE is an acceptable tool for young people with age appropriate wording and able to be completed quickly and easily. Internal consistency, based on Cronbach's alpha, showed adequate to good inter-item reliability. However, internal consistency in the non-clinical sample is slightly lower than reported by Twigg *et al.* (2016). This is largely related to item 5 ('There's been someone I felt able to ask for help' in English). Translation of this item into Italian was not easy and several versions of it were explored but the chosen translation, '*Mi sono sentito di chiedere aiuto a qualcuno quando ne ho avuto bisogno*', was considered complicated by some younger adolescents and only eventually chosen as a 'least problematic' translation when it was accepted that there seemed to be no perfect translations that would work across the en-

Table 2. Internal consistency: Cronbach's alpha values for clinical sample at pre-intervention by age band and gender.

Alpha (95% CI)	Male	Female	All
11-14	0.69 [0.50-0.80]	0.80 [0.72-0.85]	0.77 [0.70-0.82]
15-17	0.81 [0.73-0.87]	0.83 [0.79-0.86]	0.83 [0.79-0.86]
All	0.76 [0.67-0.81]	0.82 [0.79-0.85]	0.82 [0.79-0.85]

CI, confidence interval.

Table 3. Means and standard deviations by gender and age group for YP-CORE pre- and post-intervention and for change in clinical sample (n=74).

Group	No.	sd1	sd2	Mean diff.	SD diff.	CI diff.	d1	CI d1	dz	CI dz
F	48	0.70	0.54	0.77	0.57	[0.62:0.92]	1.10	[0.91:1.41]	1.37	[1.13:1.66]
M	26	0.55	0.53	0.67	0.52	[0.49:0.88]	1.22	[0.90:1.39]	1.29	[1.14:1.63]
11-14	34	0.50	0.48	0.68	0.57	[0.50:0.88]	1.36	[0.90:1.39]	1.20	[1.15:1.64]
15-17	40	0.76	0.58	0.79	0.53	[0.62:0.95]	1.04	[0.90:1.38]	1.47	[1.14:1.63]
All	74	0.67	0.55	0.74	0.55	[0.60:0.85]	1.11	[0.91:1.41]	1.35	[1.11:1.65]

sd1, standard deviation pre-intervention score; sd2, standard deviation post-intervention score; diff., difference; SD, standard deviation; CI, confidence interval; d1, effect size calculated according Cohen, 1988. Mean change divided by the standard deviation of the pre-intervention score; dz, effect size calculated according Cohen, 1988. Mean change divided by the standard deviation of the mean differences score.

tire age range. Omitting this item would produce a fractionally higher Cronbach alpha but it was retained to maintain comparability of domain coverage with the YP-CORE in other languages.

Analysis by gender and age, as in the earlier study (Twigg *et al.*, 2016) shows gender and age effects on reliability and on scores: females had higher YP-CORE scores than males, both in clinical and non-clinical samples, while older patients tended to score higher only in clinical sample. We too recommend use of gender and age specific cut-offs based on Jacobson and Truax method *c* (Jacobson & Truax, 1991) and shown in the last two rows of Table 4.

The change from baseline to termination scores in the clinical sample shows good sensitivity to change and the fairly stringent criterion for reliable change using on the baseline standard deviation of scores in this sample: 0.82. However, the 95% CI around this RCI is wide: 0.74 to 0.89. Such a stringent RCI is almost inevitable for a short, broad coverage measure. None of the 74 with both baseline and termination scores showed reliable deterioration and 35% showed reliable improvement.

The Confirmatory Factor Analysis showed perhaps surprisingly good fit to both the one-factor model proposed by O'Reilly *et al.* (2016) and the two-factor solution found by Twigg *et al.* (2009). The strong fit reflects the use of the relatively recently developed diagonally weighted least squares (DWLS) estimation instead of the traditional maximum likelihood (ML) estimation. ML es-

timation is sensitive to deviations from Gaussian distributions which are inevitable with short ordinal response scales. That the two factor model fits statistically significantly better than the one factor model is interesting. This can be interpreted as a method factor (or response set: some people answering more positively to positively cued items than negatively ones despite sharing the same general well-being as those less affected by the cueing). Alternatively, it can be interpreted as reflecting what is often thought to be genuine, if small, multidimensionality of psychological states and traits in which positive and negative aspects are strongly correlated (here $R=0.71$) but not

Table 4. Cut-off values for YP-CORE by age group and gender.

Group	CSC [95% CI]	
All	1.34 [1.27-1.41]	
11-14	1.42 [1.33-1.52]	
15-17	1.21 [1.11-1.30]	
	Female	Male
All	1.27 [1.20-1.36]	1.47 [1.36-1.59]
11-14	1.34 [1.23-1.46]	1.18 [1.07-1.30]
15-17	1.47 [1.36-1.59]	1.23 [1.07-1.37]

CSC, clinically significant change; CI, confidence interval.

Combined boxplots and trajectories for YP-CORE scores by gender
female = rose, male = blue

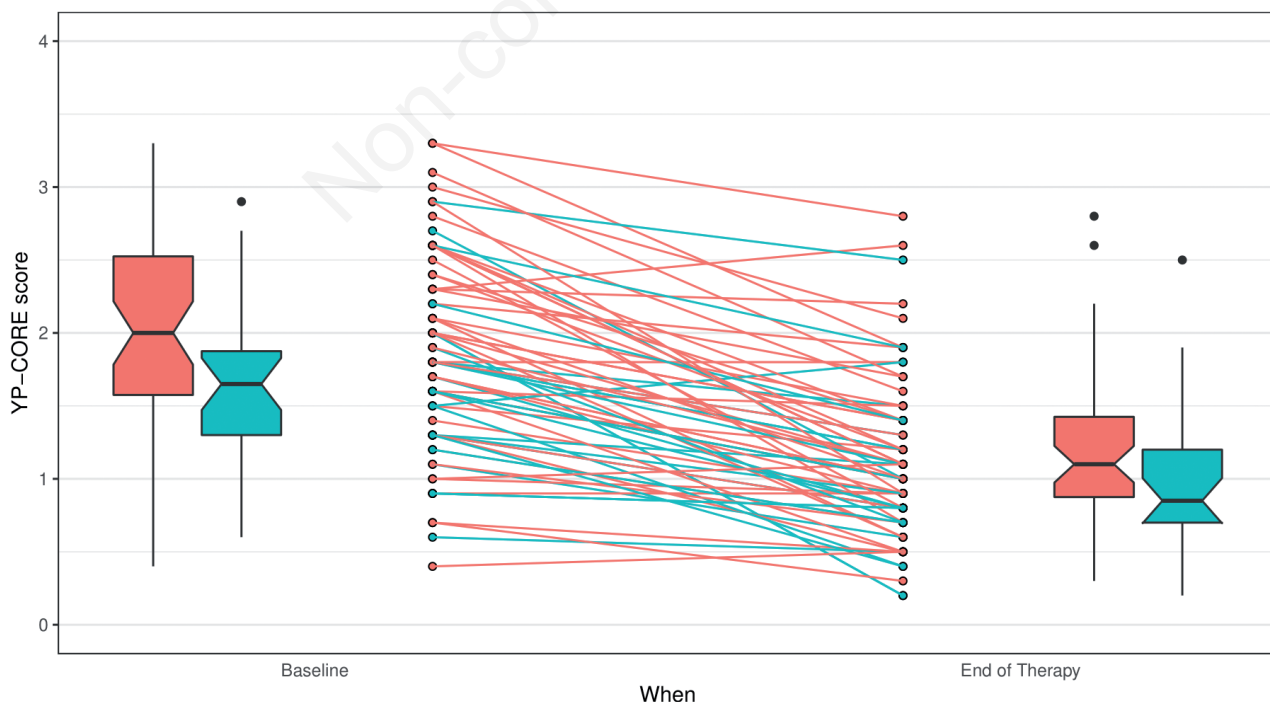


Figure 3. Boxplot and trajectories for change.

identical. The debate between the two interpretations has waged inconclusively for many years, see *e.g.* (Carmines & Zeller, 1976) for an early example arguing the effect is probably a response set *versus* most of the literature about the *positive and negative affect scales* (PANAS) for the opposite position (Watson, Clark, & Tellegen, 1988). We

are agnostic about these positions but believe that the 0.71 correlation, and the pragmatic and comparative utility of staying with a single score, justify retaining a single score across all items for the Italian YP-CORE.

The study represents a first satisfactory evidence on the psychometric quality of the Italian version of YP-CORE. However, there are limitations. The sample sizes, for both the clinical and non-clinical samples, limits the precision of the findings. The resources for the study also precluded exploring test-retest stability in either sample and convergent validity exploration against other measures (of which there are actually few or none with good translations into Italian and existing psychometric explorations). Finally, the relatively small number of subjects with post intervention scores doesn't allow conclusive estimation of the sensitivity to change in clinical contexts. These factors underline that these findings should be used cautiously and underline the values of future research with larger groups and more diverse groups of participants and exploring test-retest stability and perhaps convergent validity.

Despite these cautions we hope that clinicians will be able to introduce routine change measurement in services for young people and use their findings to explore how best to improve services, change achieved and how costs might be reduced when large datasets enable better prediction of most efficient mapping of new, and representing, clients to service options.

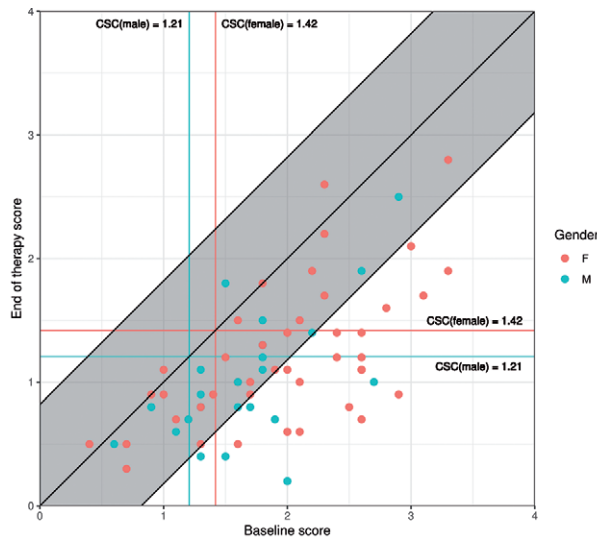


Figure 4. Jacobson plot. CSC, clinically significant change.

Table 5. Reliable clinically significant change classification of the clinical sample (n=74).

Reliable change Clinically significant change	Reliable deterioration	No reliable change No. (% of total)	Reliable improvement
Clinically deteriorated	0 (0.0%)	0 (0.0%)	-
Stayed non-clinical	0 (0.0%)	16 (21.6%)	0 (0.0%)
Stayed clinical	0 (0.0%)	14 (18.9%)	4 (5.4%)
Clinically improved	-	18 (24.3%)	22 (29.7%)

Table 6. Factor loadings of YP-CORE items at baseline.

Item	Cueing	Single factor	Two factors	
			Negative	Positive
101	Neg	0.58	0.59	
102	Neg	0.51	0.51	
103	Pos	0.37		0.48
104	Neg	0.52	0.53	
105	Pos	0.17		0.22
106	Neg	0.78	0.79	
107	Neg	0.71	0.72	
108	Neg	0.52	0.52	
109	Neg	0.79	0.79	
110	Pos	0.53		0.71
Intercorrelation factor:			0.71	

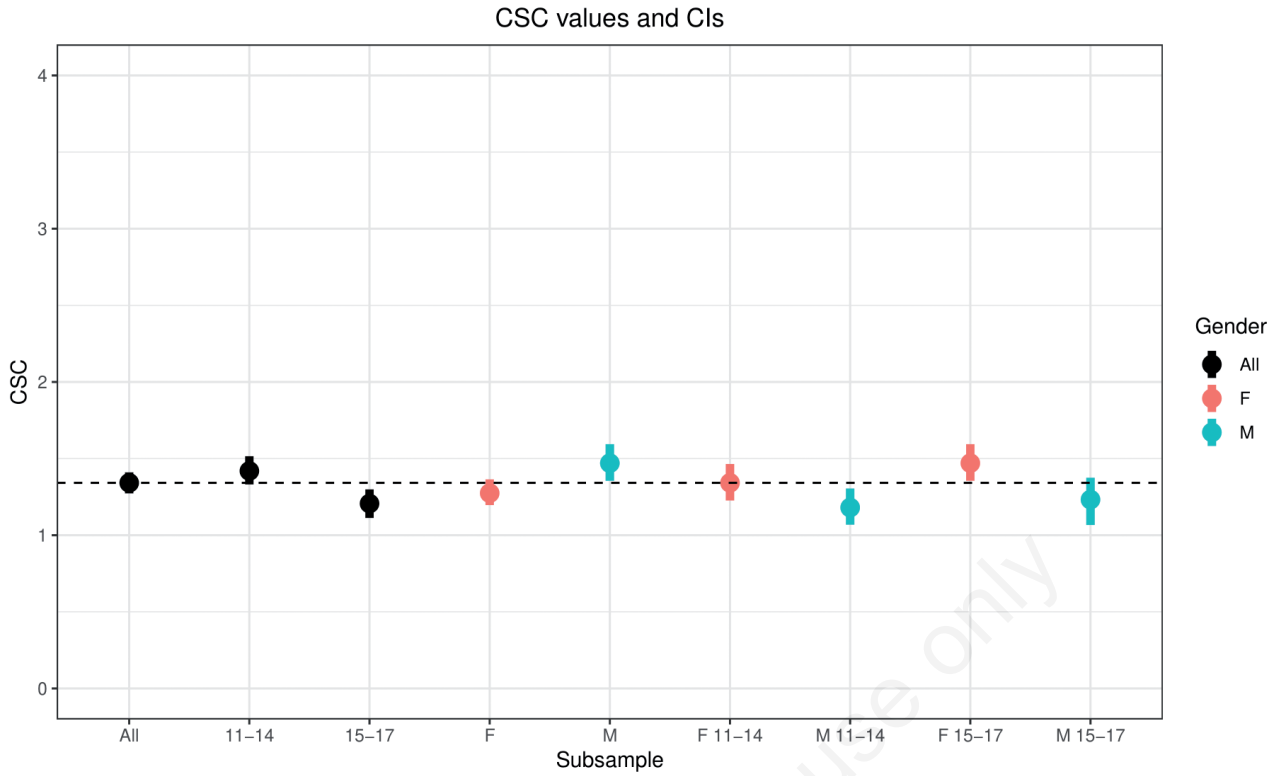


Figure 5. Clinically significant change (CSC) cutting points, entire sample and subsamples. CIs, confidence intervals.

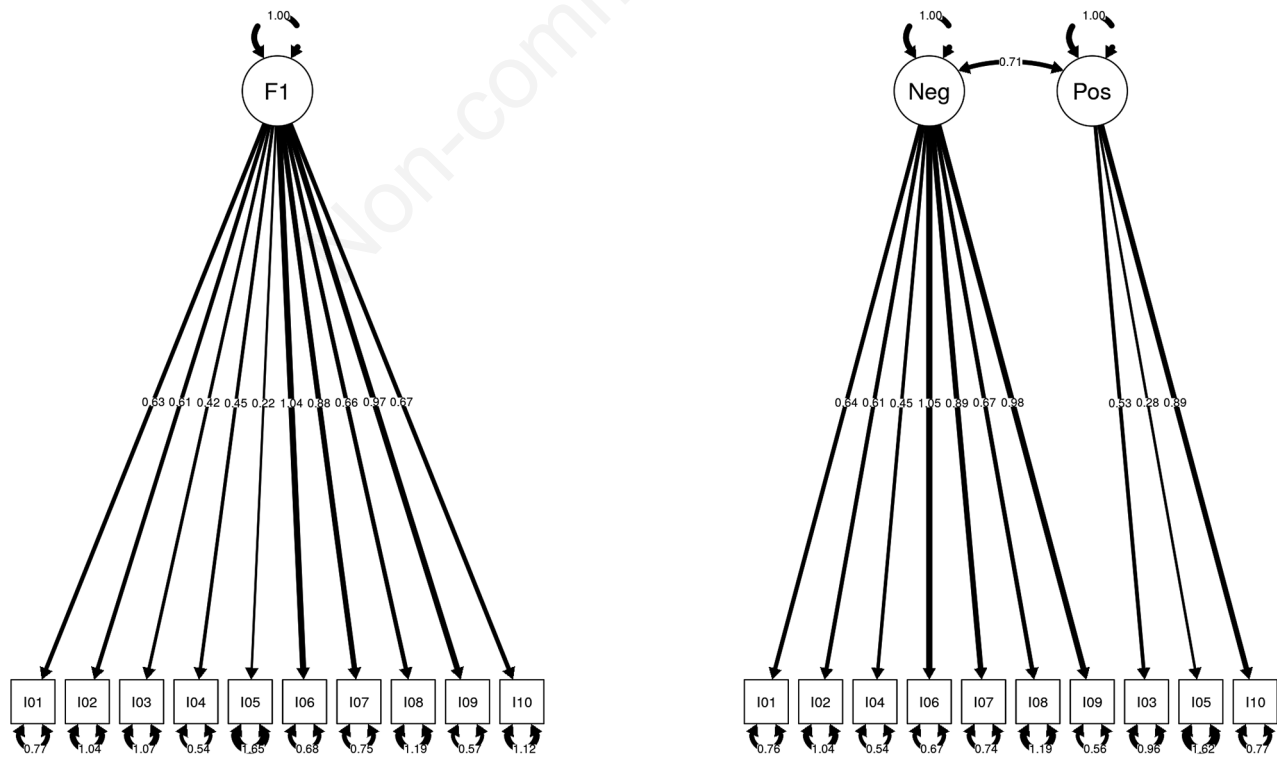


Figure 6. Structural equation models plots of single factor and two factor solutions, baseline data (n=376).

Conclusions

This study provides an initial and promising exploration of psychometric characteristics of the Italian YP-CORE. The questionnaire is copyright, so must not be changed, but is free to use without payment of any reproduction fee, is quick and easy to complete and score, seems very acceptable across diverse clinical and non-clinical young people and to professionals of various clinical orientations. It appears highly suitable for use in a wide range of psychological services and settings to detect psychological distress in Italian-speaking adolescents.

Furthermore, the progressive integration of Evidence Based Practice with Practice Based Evidence, increases the need to have reliable tools, such as the YP-CORE, to explore change across a diversity of services.

References

- Achenbach, T. M. (1991). *Integrative Guide to the 1991 CBCL/4-18, YSR, and TRF Profiles*. Burlington, VT: University of Vermont, Department of Psychology.
- Barkham, M., Margison, F., Leach, C., Lucock, M., Mellor-Clark, J., Evans, C., Benson, L., Connell, J., Audin, K., & 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*, 184-196.
- Carmines, E. G., & Zeller, R. A. (1976). *Reliability and validity assessment*. Beverly Hills: Sage Publications.
- CORE System Trust (CTS). (2015). Translation policy (internet). Available from: <https://www.coresystemtrust.org.uk/cst-translation-policy/>
- Crawford, M.J., Zoha, M., Macdonald, A.J.D., & Kingdon, D. (2017). Improving the quality of mental health services using patient outcome data: making the most of HoNOS. *BJ Psychology Bulletin, 41*(3), 172-176.
- Evans, C., Connell, J., Barkham, M., Margison, F., McGrath, G., Mellor-Clark, J., Audin, K. (2002). Towards a standardised brief outcome measure: psychometric properties and utility of the CORE-OM. *British Journal of Psychiatry, 180*, 51-60.
- Feixas, G., Badia, E., Bados, A., Medina, J. C., Grau, A., Magallón, E., Botella, L., & Evans, C. (2018). Adaptation and psychometric properties of the spanish version of the yp-core (Young Person's Clinical Outcomes in Routine Evaluation). *Actas Espanolas de Psiquiatria, 46*(3), 75-82.
- Gergov, V., Lahti, J., Marttunen, M., Lipsanen, J., Evans, C., Ranta, K., Laitila, A., & Lindberg, N. (2017). Psychometric properties of the Finnish version of the Young Person's Clinical Outcomes in Routine Evaluation (YP-CORE) questionnaire. *Nordic Journal of Psychiatry, 71*(4), 250-255.
- Goodman, R. (1997). The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry, 38*(5), 581-586.
- Gowers, S. G., Harrington, R. C., Whitton, A., Lelliott, P., Beevor, A., Wing, J., Jezzard, R. (1999a) Brief scale for measuring the outcomes of emotional and behavioural disorders in children. Health of the Nation Outcome Scales for Children and Adolescents (HoNOSCA). *British Journal of Psychiatry, 174*, 413-416.
- Gowers, S. G., Harrington, R. C., Whitton, A., Beevor, A., Lelliott, P., Jezzard, R., Wing, J. K. (1999b) Health of the Nation Outcome Scales for Children and Adolescents (HoNOSCA). Glossary for HoNOSCA score sheet. *British Journal of Psychiatry, 174*, 428-431.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 12-19.
- Kilbourne, A. M., Beck, K., Spaeth-Rublee, B., Ramanuj, P., O'Brien, R. W., Tomoyasu, N., & Pincus, H.A. (2018) Measuring and improving the quality of mental health care: a global perspective. *World Psychiatry, 17*(1), 30-38.
- Kovacs, M. (1985). The children depression inventory. *Psychopharmacology Bulletin, 21*, 995-998.
- Lavaan. (2012). Lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software, 48*(2), 1-36.
- O'Reilly, A., Peiper, N. O., Keeffe, L., Illback, R., & Clayton, R. (2016). Performance of the CORE-10 and YP-CORE measures in a sample of youth engaging with a community mental health service. *International Journal of Methods in Psychiatric Research, 25*, 324-332.
- Prakash, V., Shah, S., & Hariohm, K. (2019). Cross-cultural adaptation of patient-reported outcome measures: A solution or a problem? *Annals of Physical and Rehabilitation Medicine, 62*(3), 174-177.
- Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika, 66*(4), 507-514.
- Twigg, E., Barkham, M., Bewick, B. M., Mulhern, B., Connell, J., & Cooper, M. (2009). The Young Person's CORE: Development of a brief outcome measure for young people. *Counselling and Psychotherapy Research, September 2009, 9*(3), 160-168.
- Twigg, E., Cooper, M., Evans, C., Freire, E., Mellor-Clark, J., McInnes, B., & Barkham, M. (2016). Acceptability, reliability, referential distributions and sensitivity to change in Young Person's Clinical Outcomes in Routine Evaluation (YP-CORE) outcome measure: replication and refinement. *Child and Adolescent Mental Health, 2*, 115-123.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of a brief measures of Positive and Negative Affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*, 1063-1070.
- Yassin, S., & Evans, C. (2021). A journey to improve Arabic-speaking young peoples' access to psychological assessment tools: It's not just Google translate! *Counselling and Psychotherapy Research, capr.12431*, doi: 10.1002/capr.12431.