Quality of Life Research
https://doi.org/10.1007/s11136-024-03609-z

Cross-cultural validity and reliability of the comprehensive assessment of acceptance and commitment therapy processes (CompACT) in people with multiple sclerosis

Ambra Mara Giovannetti[1,2,3] · Rosalba Rosato[4] · Ingrid Galán[5] · Anna Toscano[4] · Elisenda Anglada[5] · Rebeca Menendez[5] · Jürgen Hoyer[6] · Paolo Confalonieri[3] · Andrea Giordano[1,4] · Kenneth Ian Pakenham[2] · Jana Pöttgen[7,8] · Alessandra Solari[1]

Accepted: 15 January 2024
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2024

Abstract

Purpose The Comprehensive assessment of Acceptance and Commitment Therapy (CompACT) is a 23-item questionnaire measuring psychological flexibility, a quality of life protective factor. An 18-item version was recently produced. We assessed validity and reliability of CompACT, and equivalence of paper and electronic (eCompACT) versions in people with multiple sclerosis (PwMS) in Italy, Germany and Spain.

Methods We used confirmatory factor analysis and assessed CompACT-23 and CompACT-18 measurement invariance between the three language versions. We assessed construct validity (Spearman's correlations) and internal consistency (Cronbach's alpha). Test–retest reliability (intraclass correlation coefficient, ICC) and equivalence of paper and eCompACT (ICC and linear regression model for repeated measures) were assessed in subsamples of PwMS.

Results A total of 725 PwMS completed the study. The three-factor structure of the CompACT-23 showed poor fit (RMSEA 0.07; CFI 0.82; SRMR 0.08), while the fit of the CompACT-18 was good (RMSEA 0.05; CFI 0.93; SRMR 0.05). Configural and partial metric invariance were confirmed, as well as partial scalar invariance (reached when five items were allowed to vary freely). The CompACT-18 showed good internal consistency (all alpha $\geq$ 0.78); and test–retest reliability (all ICCs $\geq$ 0.86). Equivalence between paper and eCompACT was excellent (all ICCs $\geq$ 0.86), with no mode, order, or interaction effects.

Conclusion Results support using the refined CompACT-18 as a three-factor measure of psychological flexibility in PwMS. Paper and eCompACT-18 versions are equivalent. CompACT-18 can be used cross-culturally, but sub-optimal scalar invariance suggests that direct comparison between the three language versions should be interpreted with caution.

Keywords Measurement invariance · CompACT · Psychological flexibility · Quality of life · Multiple sclerosis · Paper and electronic equivalence

✉ Ambra Mara Giovannetti
ambra.giovannetti@istituto-besta.it

1 Fondazione IRCCS Istituto Neurologico Carlo Besta, (Unit of Neuroepidemiology), Via Celoria 11, 20133 Milan, Italy

2 School of Psychology, University of Queensland, Brisbane, QLD, Australia

3 Fondazione IRCCS Istituto Neurologico Carlo Besta, (Unit of Neuroimmunology and Neuromuscular Diseases), Milan, Italy

4 Department of Psychology, University of Turin, Turin, Italy

5 Centre d'Esclerosi Múltiple de Catalunya—Cemcat, Barcelona, Spain

6 Technische Universitaet Dresden, Dresden, Germany

7 Institute of Neuroimmunology and Multiple Sclerosis (INIMS), Center for Molecular Neurobiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

8 Department of Neurology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

Published online: 24 February 2024

Plain English summary

Evidence suggests enhancing psychological flexibility (PF) via acceptance and commitment therapy (ACT) promotes quality of life, resilience, and mental health. A measure of PF that is validated across cultures and languages is essential for PF theory and intervention development in the international arena. A recently developed measure of PF that shows promise is the CompACT. In a prior study we translated the CompACT into three languages: Italian, German, and Spanish. In the present study we evaluate the psychometric properties of the CompACT across these different language versions in people with multiple sclerosis. As shown in other countries, our study findings show that the CompACT-18 works significantly better than the original 23-item version. The CompACT-18 has the same number of factors in all three language versions (configural invariance), and each item contributes to the latent construct to a similar extend (metric invariance), except for one item. However, comparisons of CompACT mean scores between languages should be interpreted with caution as there was little evidence of scalar invariance. The CompACT-18 demonstrated good test-retest reliability. All considered, the CompACT-18 can be used in the three different languages. The paper and electronic versions of the questionnaire are equivalent, thus offering the flexibility to use the questionnaire in different modes.

Introduction

Psychological flexibility (PF) is defined as "the ability to contact the present moment more fully as a conscious human being and, based on what the situation affords, to change or persist in behavior in order to serve valued ends" [1]. It involves six processes that can be clustered into three pillars each of which includes two therapeutic processes: (1) open (acceptance—openness to experience; cognitive defusion—observing thoughts rather than taking them literally); (2) aware (present moment awareness—mindfulness; self-as-context—contact with a sense of self that is continuous and provides flexible perspective taking); (3) engaged (values—freely chosen personally meaningful life directions; committed action—values guided effective action) [2]. The counterpart of PF is psychological inflexibility (PI), similarly composed of three pillars and six processes: (1) closed (experiential avoidance; cognitive fusion); (2) not present (disconnected from the present; attached to concepts of self); (3) unworkable actions (disconnection from values; inaction, impulsivity, or avoidance) [1].

PF skills are fostered by acceptance and commitment therapy (ACT) [3], a third-generation cognitive behavior therapy, effective in treating mental health problems in a range of contexts including chronic illness [4–6]. Evidence supports PF as the mechanism of change in ACT interventions. Mediation analyses conducted in studies on the efficacy of ACT interventions showed that increases in PF (mediator) are related to increased wellbeing [7], as well as symptom reductions in many conditions, ranging from depression and anxiety [8] to chronic pain [9].

Given the key role of PF in ACT, valid and reliable instruments that measure this construct are essential for theory and intervention development [10]. Additionally, it is important to demonstrate the theoretical integrity of ACT interventions by showing that therapeutic change occurs through PF, the mechanism of action proposed by the PF framework [11]. Accordingly, psychometrically sound measures of PF that are culturally sensitive and stable across cultures are essential [12] given that ACT interventions are implemented internationally [13–17].

PF and quality of life in chronic health conditions

Quality of life (QoL) refers to an individual's total well-being and is a multidimensional construct that has at its core the person's evaluation of his/her own life [18]. The World Health Organization defines QoL as an individual's perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns [19]. As reported in the ISOQOL Dictionary of Quality of Life and Health Outcomes Measurement, "in the context of health research, quality of life goes beyond a description of health status, but rather is a reflection of the way that people perceive and react to their health status and to other, nonmedical aspects of their lives" [20]. Hence, QoL is an important outcome variable in health-care practice and research.

PF has been demonstrated as a QoL protective factor in people with chronic health conditions [21–25]. PF in the context of chronic illness, involves activation of the openness, awareness, and engagement strategies mentioned above which enhances adaptation to health-related challenges, thereby improving QoL. In various problematic health contexts, higher PF has been correlated with lower stress, anxiety, and depression, while also being linked to better QoL [21–23, 26]. Multiple sclerosis (MS) is a chronic health condition characterized by an unpredictable course, varied and debilitating symptoms [27], and the need for ongoing adjustments in daily life [28, 29]. People with MS (PwMS) report lower QoL than people in the general population [30]. Higher PF is associated with lower self-reported distress and higher QoL in PwMS, suggesting that increasing PF could lead to a reduction in distress and improvement in QoL

[31]. Notably PF mediates the beneficial effects of resilience on distress and QoL in PwMS [24]. Moreover, ACT-based interventions in PwMS have been shown to improve resilience, QoL, and mood by strengthening PF [13, 14, 25, 32]. Therefore, increasing our understanding of the role of PF as a QoL protective factor and refining interventions that show promise in fostering PF will contribute to a holistic approach to the self-management of MS. A foundational step in pursuing these clinical and research endeavors is the validation of a reliable instrument to assess PF in PwMS across different cultures.

The assessment of PF: the comprehensive assessment of ACT processes

The Comprehensive assessment of ACT (CompACT) is a self-report patient-reported outcome measure (PROM), composed of three factors (subscales): Openness to Experience (OE) Behavioral Awareness (BA), and Valued Action (VA), which correspond to the three PF pillars, open, aware, and engaged, respectively [26]. It was developed as an alternative to the most widely used measure of PF, the Acceptance and Action Questionnaire-II (AAQ-II) that has several limitations [33–37].

We selected the CompACT because of its brevity (23 items) and demonstrated sound psychometrics in the validation study by Francis et al. [26]. The CompACT exhibited strong internal consistency and displayed patterns of convergence and divergence in line with theoretical expectations when compared to other measured factors. Specifically, elevated levels of PI correlated with increased distress, decreased health, and overall well-being. For these reasons, the authors concluded that the CompACT holds potential as an instrument for the comprehensive assessment of the ACT PF processes [26].

Despite the promising results of the original validation study [26], the CompACT's factor structure has not been replicated in subsequent studies, which were published after the commencement of the present project [12, 38–40]. A study published in 2021 showed that the CompACT performed statistically better with removal of five out of 10 items of the original Openness to Experience subscale (items 6, 13, 18, 20, and 22) all of which were negatively worded [12]. This 18-item Portuguese-adapted CompACT (CompACT-18) was tested in both Portuguese and UK samples and showed good psychometric properties [12]. Partial metric invariance was demonstrated between the two subsamples (Portuguese and UK), confirming the scale's sensitivity to contextual and cultural variations. A subsequent study evaluated both the 23- and 18-item versions of the CompACT in a sample of U.S. military personnel, and the English language CompACT-18 was found to be superior [38]. These results suggest that the factor structure and psychometric properties of the CompACT require further investigation and that a specific linguistic and psychometric validation process is needed before the CompACT can be used in different populations. Although prior studies have assessed the psychometric properties of the CompACT in different countries and languages [12, 38–40], to the best of our knowledge, no published studies have validated the CompACT in a specific clinical population.

The present study is part of an international project to develop validated versions of the CompACT in three languages. The project had two phases: (1) translation-adaptation of the CompACT into Italian, German, and Spanish [41], (2) psychometric validation of the scale for use with PwMS, in the three countries. This study constitutes phase 2, which is designed to assess the CompACT's factor structure, measurement invariance among the three language versions, psychometric properties (construct validity, internal consistency and test–retest reliability) and the equivalence of the paper and electronic (eCompACT) versions.

Materials and methods

This study was run in three European centres that belong to the "Rehabiliation in Multiple Sclerosis"—RIMS network: The Fondazione IRCCS Istituto Neurologico Carlo Besta (FINCB), Milan, Italy (coordinating centre); the University Medical Center Hamburg-Eppendorf (UKE), Hamburg, Germany; and the Centre d'Esclerosi Múltiple de Catalunya (Cemcat), Barcelona, Spain. The study was carried out in accordance with the Declaration of Helsinki recommendations. The protocol received ethical clearance from the ethics committees of the three enrolling centres: FINCB (12/09/2018, internal ref: 54; first amendment approved 12/12/2018, internal ref: 57; second amendment approved 08/05/2019, internal ref: 62), UKE (23/04/2019, clearance number: PV6040), and Cemcat (01/03/2019, clearance number: PR(AG)29/2019). All participants gave written informed consent.

Participants

A minimum target enrollment of 250 PwMS from each participating country (Italy, Germany, and Spain) was based on statistical sample size requirements and recruitment practicalities [42].

In addition to the CompACT, participants completed seven PROMs (see below). To evaluate the CompACT test–retest reliability, 150 patients completed the CompACT a second time after 10 (± 4) days. To examine the equivalence of paper and eCompact, 150 PwMS (not participating in the test–retest sub-study) completed both versions in random order (interval 10 ± 4 days).

Participants were considered eligible if the following criteria were met: neurologist confirmed MS diagnosis [43]; age ≥ 18 years; written informed consent; fluent in the language of the participating country. PwMS with severe cognitive impairment (clinical judgment) were excluded.

Outcome measures

As described previously, the 23-item CompACT measures PF and is composed of three factors (subscales): OE (10 items), BA (five items), and VA (eight items). Items are scored using a 7-point ordinal scale from 0 = strongly disagree to 6 = strongly agree, and 12 of the 23 items are reverse-coded, so that higher values always indicate greater PF [26]. CompACT raw scores were calculated by summing responses of each item in each subscale. The raw scores were then linearly transformed into scales from 0 to 100 using the following formula: $transformed\ score = \left[\frac{(raw\ score)-(lower\ possible\ raw\ score)}{possible\ raw\ score\ range}\right]\times 100$.

In addition to the translated versions of the CompACT [41], PwMS completed the following PROMs: Hospital Anxiety and Depression Scale (HADS) [44]; Connor-Davidson Resilience Scale (CDRISC)-25 [45]; 10-item AAQ-II [46]; 15-item Mindful Attention Awareness Scale (MAAS) [47]; 7-item Cognitive Fusion Questionnaire (CFQ) [48]; 20-item Valued Living Questionnaire (VLQ) [49]. Consistent with a multidimensional conceptualisation of QoL, we used the 54-item MS Quality of Life (MSQOL-54) [50] instrument to operationalise QoL. The MSQOL-54 yields two composite scores that reflect the respondent's subjective evaluation of their mental and physical QoL dimensions. These PROMs are detailed in Online Resource 1.

Clinical information and measures

Neurologists provided the following information: MS course (relapsing remitting, primary progressive, secondary progressive), time from MS diagnosis, and Expanded Disability Status Scale (EDSS) score [51]. The EDSS is the most widely used MS-specific disability scale; the EDSS score ranges between 0.0 ('normal neurological exam') and 10 ('death due to MS') [51].

Sociodemographic information

PwMS provided information on their age, gender, education, marital and working status.

Enrolment procedure

Eligible PwMS were invited to participate by neurologists at each participating centre. The centre principal investigator checked that all eligibility criteria were satisfied and then invited the consenting PwMS to complete the PROMs. Fifty PwMS in each centre were invited to complete the CompACT again after one week (test–retest sub-sample), and another 50 participants were asked to complete both paper and eCompACT in random order, one week apart (equivalence sub-sample).

In both sub-studies re-tests were performed in the participants' home (except for those preferring to complete questionnaires at the MS centre). For the equivalence sub-study, at the end of the visit, PwMS assigned to the eCompACT–paper administration order were given a closed envelope, to be opened after a week, containing the paper CompACT and a pre-addressed, return-paid envelope. One week after the test date, PwMS assigned to the paper–eCompACT order received an e-mail with the link to the website containing the eCompACT. One week after the expected completion/return date, PwMS who did not complete/return the questionnaire received a reminder (e-mail or phone call).

Statistical analyses

Continuous data were summarised using means, standard deviations, medians, and ranges/interquartile ranges, while frequencies and percentages were used for categorical data. ANOVA with t-test or Kruskall-Wallis (based on data distribution) for multiple comparisons, and chi-square tests were used to compare means or frequencies.

Factor structure and measurement invariance

As a preliminary analysis, we applied confirmatory factor analysis (CFA) with robust maximum likelihood estimator [52] to test the three-factor structure proposed by Francis et al. 2016 [26] for the CompACT-23 and the solution proposed by Trindade et al. 2021 [12] for the CompACT-18. The items removed in the CompACT-18 version (item 6, 13, 18, 20 and 22) belong to the OE subscale [12].

The best fitting model was then used to test multi-group measurement invariance between countries [53]. Three increasingly restricted levels of measurement invariance were assessed: configural (i.e., the same number of factors and the same patterns of factor loadings across groups); metric (defined as invariance of factor loadings across groups); scalar (defined as invariance of both factor loadings and item intercepts across groups). We adopted the following reference values of measurement invariance. In evaluating factor solution and configural invariance, model fit was deemed acceptable if the following criteria were met: $Chi^2/df < 3$ [54], root mean square error of approximation (RMSEA) < 0.08; comparative fit index (CFI) ≥ 0.90; and standardized root mean square residual (SRMR) ≤ 0.08 [55, 56]. Akaike Information Criterion (AIC) and Bayesian

Information Criterion (BIC) were used to compare models, with lower values indicating better model fit.

A worsening of CFI, exceeding the threshold of 0.01 and accompanied by a change in RMSEA ≥ 0.015 or a change in SRMR of ≥ 0.03, was considered to indicate no metric invariance. When considering scalar invariance, the cutoff values for change in CFI and RMSEA were the same as for metric invariance, whereas a cutoff of 0.01 was considered for SRMR [57].

Reliability and validity

Internal consistency was assessed using Cronbach's alpha (benchmark value > 0.70) [58]. Test–retest reliability was assessed with the intraclass correlation coefficient (ICC) with 95% confidence intervals (CIs; benchmark value > 0.75) [59].

Spearman's correlation coefficients were used to investigate convergent validity. Consistent with the PF/PI framework, based on previous studies [12, 26] we expected the CompACT to have moderate to strong positive correlations with the PF processes (present moment awareness, values-consistent behavior), and moderate to strong negative correlations with the PI processes (experiential avoidance and cognitive fusion), moderate positive correlation with resilience and QoL, and strong negative correlation with distress (anxiety and depression).

Equivalence of electronic and paper versions

The equivalence of the eCompACT and paper versions was assessed using a randomized cross-over design, which allows for testing of administration mode (paper [P] vs. electronic [E]) and order (P-E vs. E-P) effects, and their interaction. A linear regression model for repeated measures design, with an exchangeable correlation structure of the errors, was applied using the following design variables: administration mode, administration order, and their interaction; all estimates were adjusted for language (Italian, Spanish, Germany), gender, age and EDSS (median values as cutoff). In assessing equivalence, model fit was evaluated using likelihood ratio test.

All analyses were performed using SPSS 28, except for measurement invariance for which Mplus 8 was used. All statistical tests were two-tailed, and considered significant at $p < 0.05$.

Results

Sample characteristics

Between April 2019 and May 2020, a total of 725 PwMS completed the study (Italy 251; Germany 226; Spain 248; Table 1). PwMS mean age was 45.2 (SD 11.7). Most

Table 1 Demographic and clinical characteristics for the total sample and country sub-samples

	Total (N = 725) Mean (SD)	Italy (N = 251)	Germany (N = 226)	Spain (N = 248)	P value
Age (years)	45.2 (11.7)	41.9 (10.0)$_{a,b}$	44.5 (12.3)$_{b,c}$	49.4 (11.6)$_{a,c}$	< 0.001
Women, n (%)	487 (67.2)	191 (76.1)$_{a,b}$	139 (61.5)$_{a}$	157 (63.3)$_{b}$	0.001
Marital status, n (%)					
Single	183 (25.3)	97 (38.6)	46 (20.3)	40 (16.2)	
Married	476 (65.7)	138 (55.0)	167 (73.9)	171 (69.2)	
Divorced	47 (6.5)	14 (5.6)	11 (4.9)	22 (8.9)	
Widow	18 (2.5)	2 (0.8)	2 (0.9)	14 (5.7)	< 0.001
Education, n (%)					
Primary school	46 (6.4)	1 (0.4)	14 (6.2)	31 (12.5)	
Middle school diploma	102 (14.1)	15 (6.0)	55 (24.3)	32 (12.9)	
High school diploma	222 (30.6)	103 (41.0)	37 (16.4)	82 (33.1)	
Degree	306 (42.2)	97 (38.7)	110 (48.7)	99 (39.9)	
PhD/Specialization	49 (6.8)	35 (13.9)	10 (4.4)	4 (1.6)	< 0.001
Disease duration (years)	11.9 (9.1)	10.5 (7.8)	8.6 (7.5)	16.5 (9.8)	< 0.001
MS type, n (%)					
Relapsing remitting	530 (73.2)	231 (92.0)	154 (68.4)	145 (58.5)	
Secondary progressive	114 (15.7)	11 (4.4)	30 (13.3)	73 (29.4)	
Primary progressive	80 (11.1)	9 (3.6)	41 (18.3)	30 (12.1)	< 0.001
EDSS score, median (min–max)	2.5 (0.0–8.5)	2.0 (0.0–8.5)	2.5 (0.0–8.0)	4.0 (0.0–8.5)	< 0.001

SD standard deviation. *EDSS* Expanded Disability Status Scale

participants were women (67%), had high-school diploma or higher (80%), and lived with a partner (66%). Mean MS duration was 12 years (SD = 9.1). Most participants (73.2%) had relapsing MS; the median EDSS score was 2.5 and ranged between 0.0 ('normal neurological exam') and 8.5 ('essentially restricted to bed much of day').

Distribution of the PROMs in addition to the CompACT, and comparison across country sub-samples are reported in Table 2.

Factor structure

Before testing the factor structure of the CompACT, we added two residual correlations between items 16 (*"I do jobs or tasks automatically, without being aware of what I'm doing"*) and 19 (*"It seems I am 'running on automatic' without much awareness of what I'm doing"*) and items 10 (*"I behave in line with my personal values"*) and 21 (*"My values are really reflected in my behavior"*) as these item pairs were worded similarly and the covariance between them cannot be explained solely by their relationship with the latent construct. Then the original three-factor model of the CompACT-23 was tested on the total sample. The results showed poor statistical fit (Chi^2/df, 989.9/225 = 4.40; RMSEA = 0.068; CFI = 0.819; SRMR = 0.084). As next step, the three-factor model of the CompACT-18 [12] was tested on the total sample. All fit indices indicated an acceptable fit (Chi2/df 338.4/130 = 2.60; RMSEA 0.047; CFI 0.935; SRMR 0.051) and both AIC and BIC decreased compared with the previous model (Table 3).

Table 2 Distribution of the other patient-reported outcome measure scores, and comparison across the country sub-samples

	Total (N = 725)		Italy (N = 251)		Germany (N = 226)		Spain (N = 248)		*P value*
	Mean (SD)	Median (Min–Max)	Mean (SD)	Median (Min–Max)	Mean (SD)	Median (Min–Max)	Mean (SD)	Median (Min–Max)	
MSQOL-54 MHC	66.4 (18.2)	71.4 (8.4–92.9)	68.5 (16.2)	$72.9_{a,b}$ (26.0–91.9)	64.4 (20.6)	$72.8_{a,c}$ (8.4–91.9)	66.2 (17.7)	$70.2_{b,c}$ (19.7–92.9)	< 0.001
MSQOL-54 PHC	63.0 (18.6)	64.1 (16.4–98.1)	70.4 (18.5)	74.8 (18.2–98.1)	62.8 (17.2)	65.3 (17.4–89.8)	55.8 (17.0)	53.5 (16.4–88.5)	0.231
HADS-A	8.4 (4.4)	8.0 (0.0–20.0)	6.6 (4.0)	6.0_{a} (0.0–19.0)	6.8 (4.6)	6.0_{b} (0.0–20.0)	11.6 (2.4)	$12.0_{a,b}$ (5.0–16.0)	< 0.001
HADS-D	5.9 (3.8)	6.0 (0.0–21.0)	3.8 (3.2)	$3.0_{a,b}$ (0.0–14.0)	5.2 (4.3)	$4.0_{a,c}$ (0.0–21.0)	8.6 (1.7)	$9.0_{b,c}$ (4.0–15.0)	< 0.001
CD-RISC	65.5 (16.2)	68.0 (15.0–100.0)	61.5 (17.1)	$62.0_{a,b}$ (21.0–100.0)	68.8 (15.4)	71.0_{a} (15.0–97.0)	66.7 (15.1)	68.0_{b} (27.0–99.0)	< 0.001
AAQ-II	28.1 (10.3)	25.5 (10.0–66.0)	29.5 (10.3)	27.0_{a} (10.0–59.0)	23.8 (5.1)	$22.0_{a,b}$ (13.0–42.0)	30.4 (12.3)	28.0_{b} (10.0–66.0)	< 0.001
MAAS	4.4 (0.9)	4.5 (1.5–6.0)	4.3 (0.8)	4.5 (2.0–6.0)	4.3 (1.0)	4.5 (1.5–5.9)	4.4 (1.0)	4.5 (2.0–6.0)	0.460
CFQ	21.6 (9.5)	21.0 (7.0–49.0)	22.2 (8.5)	22.0_{a} (7.0–49.0)	20.3 (10.6)	$17.0_{a,b}$ (7.0–49.0)	22.1 (9.3)	22.0_{b} (7.0–49.0)	0.007
VLQ	55.9 (18.0)	61.8 (1.0–100.0)	60.9 (18.3)	67.0_{a} (1.0–100.0)	49.3 (17.0)	$55.1_{a,b}$ (3.5–98.0)	57.0 (16.8)	64.7_{b} (9.3–100.0)	< .001

MSQOL-54 MHC, 54-item Multiple Sclerosis Quality of Life Mental Health Composite; MSQOL-54 PHC, MS-54 QOL Physical Health Composite; HADS-A is Hospital Anxiety and Depression Scale – Anxiety, HADS-D, HADS– Depression; CD-RISC, Connor-Davidson Resilience Scale; AAQ, Acceptance and Action Questionnaire; MAAS = Mindful Attention Awareness Scale; CFQ, Cognitive Fusion Questionnaire; VLQ, Valued Living Questionnaire. SD, standard deviation. Medians having the same subscript (a, b, c) are significantly different from each other in post hoc multiple comparisons (Kruskall-Wallis test)

Table 3 Goodness-of-fit indices for CFA models of CompACT-23 and CompACT-18

Model		Number free parameters	AIC	BIC	$Chi^{2°}$/df	RMSEA	CFI	SRMR
23-item	N = 725	74	60,387.2	60,726.5	989.9/225	0.068	0.819	0.084
18-Item	N = 725	59	46,251.0	46,521.6	338.4/130	0.047	0.935	0.051

AIC Akaike information criterion; *BIC* Bayesian information criterion; *DF* Error degree of freedom; *RMSEA* root mean square error of approximation; *CFI* comparative ft index; *SRMR* standardized root mean square residual. ° p value < .001. Both models account for the error correlation between items 16 and 19 and items 10 and 21

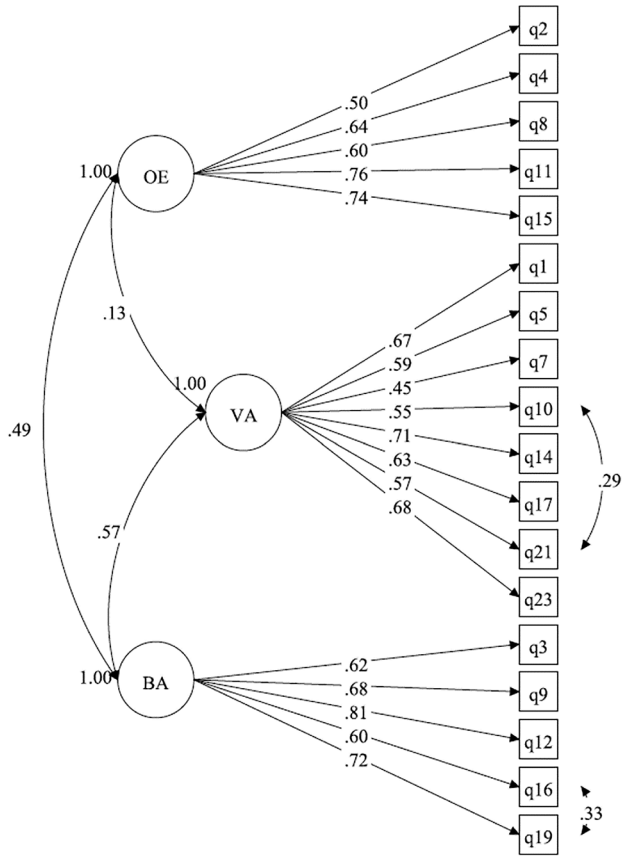

Fig. 1 Loadings of 18 items for total sample. Factor structure of the CompACT-18 items. *OE* Openness to Experience, *VA* Valued Action; *BA* Behavioural Awareness. All estimates are statistically significant $p < 0.01$

Figure 1 shows the estimated factor loadings for the 18-item version in the whole sample. All estimates are large and statistically significant.

Measurement invariance between the Italian, German, and Spanish CompACT versions

To assess the measurement invariance of the CompACT-18 scale across the three language versions, we first tested the 3-factor model in each country sub-sample (Table 3, M1, M2, and M3; loadings of 18 items for each sub-sample are reported in Online Resource 2). In these analyses, the ratio of chi-squared to degree of freedom is < 3 and all fit indices are within the bounds, except for CFI for the Italian and German samples, which is slightly below threshold 0.90.

The model fitted the data well in terms of RMSEA and SRMR, and CFI was slightly under the cut-off indicating that configural invariance is substantially achieved (Table 4, M4: RMSEA 0.062, CFI 0.896, SRMR 0.068), whereas it did not support metric invariance as indicated by ΔCFI > 0.01 (Table 4, M5: ΔRMSEA 0.001; ΔCFI 0.011; ΔSRMR 0.013), suggesting that the corresponding latent factors might have different significance in the different groups. Therefore, we tested whether partial measurement invariance was achieved. In the model in which the loading for item 4 ('*I try to stay busy to keep thoughts or feelings from coming'*) was allowed to vary freely between samples, the results showed good model fit: ΔRMSEA 0.000; ΔCFI 0.009; ΔSRMR 0.012. Thus, the remaining 17 items presented factor invariance and contributed to their respective subscales in very similar ways across translations and cultures. We then tested scalar invariance (i.e., whether item intercepts were also invariant across groups). The

Table 4 CompACT-18 measurement invariance across language versions

Models		Number free parameters	AIC	BIC	Chi$^{2°}$/df	Goodness-of-fit			Comparison		
						RMSEA	CFI	SRMR	ΔRMSEA	ΔCFI	ΔSRMR
	Languages										
M1	German (N = 226)	59	14,630.0	14,831.8	250.4/130	.064	.898	.072			
M2	Italian (N = 251)	59	15,513.8	15,721.8	282.2/130	.068	.887	.066			
M3	Spanish (N = 248)	59	15,740.8	15,948.1	218.8/130	.052	.907	.064			
	Multigroup invariance										
M4	Configural invariance	177	45,884.6	46,696.4	751.2/390	.062	.896	.068			
M5	Metric invariance	147	45,910.0	46,584.2	818.3/420	.063	.885	.081	0.001	0.011	0.013
M5^	Partial metric invariance	149	45,905.4	46,588.8	811.2/418	.062	.887	.080	< 0.001	0.009	0.012
M6	Scalar invariance	119	46,070.8	46,616.5	1012.8/448	.072	.838	.090	0.01	0.049	0.01
M6*	Partial scalar invariance	129	45,922.1	46,513.7	864.4/438	.063	.877	.083	0.001	0.01	0.003

AIC Akaike information criterion; *BIC* Bayesian information criterion; *DF* Error degree of freedom; *RMSEA* root mean square error of approximation; *CFI* comparative ft index; *SRMR* standardized root mean square residual. °p value < .001. All models account for error correlation between items 16/19 and items 10/21. ^ loading for item 4 'I try to stay busy to keep thoughts or feelings from coming ' freely estimated; *Intercept for items [Q7, Q2, Q10, Q8, Q19] freely estimated

deterioration of the CFI index from the metric to the scalar model indicates that the model is not invariant across groups in terms of item intercepts (Table 4, M6: ΔRMSEA 0.010; ΔCFI 0.049; ΔSRMR 0.010), which means that the mean items are not the same across the three groups. Five items contributed to this result: (Item 2: *'One of my big goals is to be free from painful emotions'*; Item 7: *'I make choices based on what is important to me. even if it is stressful'*, Item 8: *'I tell myself that I shouldn't have certain thoughts'*, Item 10: *'I behave in line with my personal values'*; Item 19: *'It seems I am "running on automatic" without much awareness of what I'm doing'*). The model in which these five items were allowed to vary freely showed good model fit: (Table 4, M6*: ΔRMSEA 0.001; ΔCFI 0.010; ΔSRMR 0.003). These results indicate that the CompACT-18 exhibits partial metric and partial scalar invariance.

Descriptive values of the CompACT subscales for the whole sample and each subsample are shown in Table 5.

Reliability and validity

Internal consistency and validity were reported for the CompACT-18. Cronbach's alphas for the CompACT-18 subscales for the total sample were > 0.70 (OE, α 0.78; BA, α 0.83; VA, α 0.82).

Test–retest reliability of the CompACT-18 calculated on 153 PwMS (Germany N = 53, Spain N = 50, Italy N = 50) was excellent for each subscale: OE, ICC 0.87 (95% CI 0.82–0.91); BA, ICC 0.86 (95% CI 0.80–0.90); VA, ICC 0.88 (95% CI 0.82–0.91).

The inter-correlations between the CompACT-18 subscales were positive and of moderate magnitude, except for the correlation between OE and VA scales, which was negligible: OE–BA, ρ 0.43 (95% CI 0.36–0.48); OE–VA, ρ 0.12 (95% CI 0.04–0.19); BA–VA, ρ 0.44 (95% CI 0.38–0.50).

The CompACT-18 subscales showed good convergent validity. As expected, CompACT-18 subscales were moderately and negatively correlated with experiential avoidance (AAQ-II, ρ −0.49 to −0.43) and cognitive fusion (CFQ, ρ −0.52 to −0.43), and positively correlated with mindfulness (MAAS, ρ 0.66 to 0.29). Valued living scores (i.e., the importance of certain life domains and the consistency of behaviors; VLQ) were not significantly correlated with the OE subscale and the correlations between valued living scores and the other CompACT-18 subscales were moderate (ρ 0.35 to 0.22). In addition, the three CompACT-18 subscales were also moderately, negatively associated with anxiety (HADS-A ρ -0.35 to −0.30) and depression (HADS-D ρ −0.45 to −0.31). Regarding QoL, higher scores on the CompACT-18 subscales BA and VA were significantly related to better mental and physical QoL. The same CompACT-18 subscales were also positively correlated with resilience (CD-RISC 25). Data on convergent validity are reported in Table 6.

Equivalence of paper and electronic versions

A total of 136 PwMS participated in the equivalence substudy (Germany N = 36, Spain N = 49 and Italy N = 51). Administration order was paper/eCompACT (N = 77), and eCompACT/paper (N = 59).

Reliability was excellent for each subscale: OE, ICC 0.87 (95% CI 0.81–0.91); BA, ICC 0.86 (95% CI 0.80–0.90); VA, ICC 0.88 (95% CI 0.82–0.91).

The linear regression model for repeated measures showed that all mode and order effects as well as their interaction effects were statistically non-significant (all *p* values > 0.5). These results confirmed the equivalence of the two versions [Online Resource 3].

Discussion

The aim of this study was to validate the CompACT cross-culturally for its use with PwMS who face life-changing difficulties related to the progressive yet unpredictable nature

Table 5 Means and standard deviations for CompACT-18 subscale scores for the total sample and by country

	Total (N = 725)		Italy (N = 251)		Germany (N = 226)		Spain (N = 248)		P value
	Mean (SD)	Median (Min–Max)	Mean (SD)	Median (Min–Max)	Mean (SD)	Median (Min–Max)	Mean (SD)	Median (Min–Max)	
CompACT-18 subscale									
Openness to experience	44.6 (24.2)	43.3 (0.0–100.0)	47.8 (26.2)	43.3_{a} (0.0–100.0)	48.8 (23.9)	50.0_{b} (0.0–100.0)	37.6 (20.8)	$33.3_{a,b}$ (0.0–96.7)	< 0.001
Behavioral awareness	65.3 (24.8)	70.0 (18.8–100.0)	71.6 (24.6)	76.7_{a} (0.0–100.0)	63.7 (25.1)	66.7_{a} (6.7–100.0)	60.2 (23.3)	60.0_{a} (6.7–100.0)	< 0.001
Valued Action	79.7 (15.0)	83.3 (18.8–100.0)	84.0 (14.2)	87.5_{a} (29.2–100.0)	78.6 (15.3)	81.2_{a} (18.7–100.0)	76.4 (14.6)	79.2_{a} (29.2–100.0)	< 0.001

Medians having the same subscript (a,b,c) are significantly different from each other in post hoc multiple comparisons (Kruskall-Wallis test)

Table 6 Spearman correlation (ρ) between CompACT-18 subscale scores and concurrent validity measures for the total sample N = 725

	Openness to experience		Behavioral awareness		Valued action	
	ρ	95% CI	ρ	95% CI	ρ	95% CI
MSQOL-MHC	0.39	0.33; 0.45	0.56	0.51; 0.61	0.51	0.46; 0.56
MSQOL-PHC	0.22	0.15; 0.29	0.36	0.29; 0.24	0.42	0.36; 0.48
HADS-A	−0.33	−0.39; −0.26	−0.35	−0.41; 0.28	−0.30	−0.37; −0.24
HADS-D	−0.31	−0.38; −0.25	−0.44	−0.50; 0.38	−0.45	−0.51; −0.39
CD-RISC	0.16	0.08; 0.23	0.38	0.32; 0.44	0.58	0.53; 0.62
AAQ-II	−0.43	−0.49; −0.37	−0.49	−0.54; −0.43	−0.46	−0.51; −0.40
MAAS	0.29	0.22; 0.36	0.66	0.62; 0.70	0.39	0.32; 0.45
CFQ	−0.49	−0.55; −0.44	−0.52	−0.58; 0.47	−0.43	−0.49; −0.37
VLQ	0.00*	−0.07; 0.08	0.22	0.15; 0.29	0.35	0.28; 0.41

95% CI = 95% Confidence Interval. MSQOL-54 MHC, 54-item Multiple Sclerosis Quality of Life Mental Health Composite; MSQOL-54 PHC, MS-54 QOL Physical Health Composite; HADS-A is Hospital Anxiety and Depression Scale – Anxiety, HADS-D, HADS– Depression; CD-RISC, Connor-Davidson Resilience Scale; AAQ, Acceptance and Action Questionnaire; MAAS = Mindful Attention Awareness Scale; CFQ, Cognitive Fusion Questionnaire; VLQ, Valued Living Questionnaire. All p values are < 0.001 except for* > .05

of MS. As reported above, PF has been empirically substantiated as a protective determinant of QoL among PwMS, as evidenced by extant research [13, 14, 24, 25, 31, 32]. Consequently, to better understand the role of PF in enhancing QoL it is imperative that psychometrically sound and culturally sensitive measures of PF be established. This, in turn, will also facilitate the nuanced refinement of PF-based interventions designed to foster QoL in PwMS.

In line with previous studies [12, 38, 39], the three-factor structure of the CompACT-23 showed poor fit with item 6 loading on a different factor than in the original solution, and some items (i.e., 13, 18, 20, 22) had cross-loadings. These results are consistent with the two studies in which the original factor structure of the questionnaire was tested in Portuguese and U.S. military personnel samples, respectively [12, 38]. In both studies, a refined version of the CompACT (18-item version) performed statistically better than the original one [12, 38], indicating that the original structure has some weaknesses. Item pruning is supported not only by psychometric results but also by two design issues. As reported by Hsu et al. [39], two aspects may have played a role in the misfit of the CompACT-23 model. First, the redundancy in content between items (i.e., items 13 and 22). Second, the presence of either positive or negative wording across or within the factors. While the BA subscale consists of negatively worded items and VA of positively worded items, the OE items are mixed (three out of ten are positively worded). By eliminating items with redundant content and inconsistent wording direction within a factor, the CompACT-18 represents an improved version of the original questionnaire.

The configural invariance and partial metric invariance of the CompACT-18 were confirmed in our study. This means that the CompACT has the same number of factors in all the three language versions (configural invariance), and that each item contributes to the latent construct to a similar extent (metric invariance), except for one item (item 4). Partial scalar invariance was also achieved but only when five items were allowed to vary freely. This suggests that comparison of mean CompACT scores between languages should be interpreted with caution. Further studies are necessary to modify, integrate, or eliminate specific items that render the instrument non-fully invariant cross-culturally. Chen [57] demonstrated that as the proportion of non-invariant items on a factor increased, the bias of the mean estimates for the subgroups (and thus the estimated difference between the means of the subgroups) also increased but gave no indication of an "acceptable" proportion of invariant items. A Monte-Carlo simulation showed that metric non-invariance (unequal factor loadings) has a negligible effect on the mean differences of a latent factor, but that scalar non-invariance (unequal intercepts) leads to serious misinterpretation of the true mean differences [60, 61]. For all these reasons, further studies should investigate the scalar invariance of the 18-item CompACT. An important step would be to assess whether some items should be linguistically adjusted or deleted to achieve more robust scalar invariance across Italian, German, and Spanish languages.

The CompACT-18 showed good internal consistency, with Cronbach's alpha values for the subscales ranging from good (OE, 0.78) to acceptable (BA 0.83, VA 0.82). These results are consistent with those from studies of different CompACT versions (i.e., 23-item [26], 18-item [12, 38], and 15-item [39]). Our study confirmed the stability of CompACT-18 over time (ICCs 0.86 or above). Consistent with our hypotheses and the literature [12, 25, 38, 39], the CompACT-18 subscales showed adequate convergent validity in the expected directions with measures of PI (experiential avoidance and cognitive fusion) and PF (mindfulness and

valued living), as well as with other related constructs (i.e., anxiety, depression, QOL and resilience). Given the importance of QoL in the context of chronic health conditions, it is noteworthy that the significant correlation between higher PF and better QoL supports the role of PF as a QoL protective factor and underscores the potential usefulness of the CompACT in QoL research.

Availability of ePROMs has advantages compared to paper version, including easy integration with electronic health records and disease registries, less missing data, reduction of computational burden and errors, and (for adaptive versions) reduced administration burden. However, it is mandatory to formally assess the equivalence and acceptability of different modes of administration to have confidence in the validity of data collected with different modes of administration and to allow pooled analyses when different versions are used within and between studies [62–64]. This was the first study to examine the equivalence of the paper and eCompACT versions. The results show that the two modes of administration are equivalent.

This study has the following limitations. First, generalizability of findings is limited given they were derived from a sample of PwMS. Further research should investigate the psychometrics of the CompACT-18 using community samples in each country. Second, PwMS were recruited from one centre in each country, which may have limited the representativeness of PwMS. However, each MS centre involved in the study is a national MS centre of excellence and services patients from the entire country, and the clinical characteristics of the sample (e.g. EDSS score, MS type) were well varied. Third, participants were required to answer all items before saving responses to the e-survey, therefore it was not possible to assess the rate of missing items, an important aspect to consider when assessing questionnaire feasibility. However, no PwMS dropped out from the study. Finally, further studies should assess the predictive validity and responsiveness of the CompACT-18 [65].

Conclusion

The results of this multicentre study supported the psychometric properties and measurement invariance of CompACT-18 in Italian, German and Spanish when administered to PwMS. The results suggest that the CompACT-18 can be used in the three language versions, although direct score comparisons between languages should be interpreted with caution. The CompACT-18 has excellent reliability and good convergent validity. Our findings also show that the paper and electronic versions of the CompACT-18 are equivalent. Having a valid and reliable measure of PF for use with PwMS is particularly important, as a recent study demonstrated that greater PF is related to better coping outcomes (lower distress, higher QoL) in this population [31].

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11136-024-03609-z.

Acknowledgements Authors want to thank the authors of the original version of the CompACT (Francis AW, Dawson DL, Golijani-Moghaddam N) for their support and useful insights during all project phases, and all the PwMS who participated in the study.

Author contributions AMG, AS and KIP conceived and developed the study protocol. AS oversaw the study, AMG was the principal investigator of the study and of the Italian centre, JP was the principal investigator of the German centre, IG was the principal investigator of the Spanish centre, EA and RM were responsible for data collection in Spain. JP developed the online survey. RR and AT performed the statistical analyses. AMG, RR and AS drafted the manuscript. All authors approved and edited the final manuscript.

Funding This study is supported by the "REHABILITATION IN MULTIPLE SCLEROSIS -RIMS European network for best practice and resource" (RIMS GRANT PROGRAM 2018 to AMG). The translation of the CompACT into italian was supported by FISM - Fondazione Italiana Sclerosi Multipla – cod. 2016/B/3 and financed or co-financed with the '5 per mille' public funding. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data availability Data are available upon request to the corresponding author.

Declarations

Competing Interest AS reports grants from Fondazione Italiana Sclerosi Multipla (FISM) while conducting the study; AS reports personal fees from Biogen Idec, Merck Serono, Novartis, Almirall, and Excemed. JP reports grants from Merck Serono and BMS Celgene.

Ethical approval This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Ethics Committees of the three enrolling centres: FINCB (12/09/2018, internal ref: 54; first amendment approved 12/12/2018, internal ref: 57; second amendment approved 08/05/2019, internal ref: 62), UKE (23/04/2019, clearance number: PV6040), and Cemcat (01/03/2019, clearance number: PR(AG)29/2019).

Consent to participate Informed consent was obtained from all individual participants included in the study.

References

1. Luoma, J. B., Hayes, S. C., & Walser, R. D. (2007). *Learning ACT: An acceptance & commitment therapy skills-training manual for therapists*. New Harbinger & Reno.
2. Luoma, J. B., Hayes, S. C., & Walser, R. D. (2017). *Learning ACT: An acceptance & commitment therapy skills-training manual for therapists* (2nd ed.). Oakland: New Harbinger & Reno.
3. Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (1999). *Acceptance and commitment therapy: An experiential approach to behavior change*. Guilford Press.

4. Brandon, S., Pallotti, C., & Jog, M. (2021). Exploratory study of common changes in client behaviors following routine psychotherapy: Does psychological flexibility typically change and predict outcomes? *Journal of Contemporary Psychotherapy, 51*(1), 49–56. https://doi.org/10.1007/s10879-020-09468-2
5. Gloster, A. T., Walder, N., Levin, M. E., Twohig, M. P., & Karekla, M. (2020). The empirical status of acceptance and commitment therapy: A review of meta-analyses. *Journal of Contextual Behavioral Science, 18*, 181–192. https://doi.org/10.1016/j.jcbs.2020.09.009
6. Hayes, S. C., Levin, M. E., Plumb-Vilardaga, J., Villatte, J. L., & Pistorello, J. (2013). Acceptance and commitment therapy and contextual behavioral science: Examining the progress of a distinctive model of behavioral and cognitive therapy. *Behavior therapy, 44*(2), 180–198. https://doi.org/10.1016/j.beth.2009.08.002
7. Wersebe, H., Lieb, R., Meyer, A. H., Hofer, P., & Gloster, A. T. (2018). The link between stress, well-being, and psychological flexibility during an acceptance and commitment therapy self-help intervention. *International Journal of Clinical and Health Psychology, 18*(1), 60–68. https://doi.org/10.1016/j.ijchp.2017.09.002
8. Fledderus, M., Bohlmeijer, E. T., Fox, J.-P., Schreurs, K. M. G., & Spinhoven, P. (2013). The role of psychological flexibility in a selfhelp acceptance and commitment therapy intervention for psychological distress in a randomized controlled trial. *Behaviour Research and Therapy, 51*(3), 142–151. https://doi.org/10.1016/j.brat.2012.11.007
9. Hughes, L. S., Clark, J., Colclough, J. A., Dale, E., & McMillan, D. (2017). Acceptance and commitment therapy (ACT) for chronic pain: A systematic review and meta-analyses. *The Clinical Journal of Pain, 33*(6), 552–568. https://doi.org/10.1097/AJP.0000000000000425
10. Subhas, N., Pang, N. T., Chua, W. C., Kamu, A., Ho, C. M., David, I. S., Goh, W. W., Gunasegaran, Y. I., & Tan, K. A. (2021). The cross-sectional relations of COVID-19 fear and stress to psychological distress among frontline healthcare workers in Selangor, Malaysia. *International Journal of Environmental Research and Public Health, 18*(19), 10182. https://doi.org/10.3390/ijerph181910182
11. Stockton, D., Kellett, S., Berrios, R., Sirois, F., Wilkinson, N., & Miles, G. (2019). Identifying the underlying mechanisms of change during acceptance and commitment therapy (ACT): A systematic review of contemporary mediation studies. *Behavioural and Cognitive Psychotherapy, 47*(3), 332–362. https://doi.org/10.1017/S1352465818000553
12. Trindade, I. A., Ferreira, N. B., Mendes, A. L., Ferreira, C., Dawson, D., & Golijani-Moghaddam, N. (2021). Comprehensive assessment of acceptance and commitment therapy processes (CompACT): Measure refinement and study of measurement invariance across Portuguese and UK samples. *Journal of Contextual Behavioral Science, 21*, 30–36. https://doi.org/10.1016/j.jcbs.2021.05.002
13. Pakenham, K. I., Mawdsley, M., Brown, F. L., & Burton, N. W. (2018). Pilot evaluation of a resilience training program for people with multiple sclerosis. *Rehabilitation Psychology, 63*(1), 29–42. https://doi.org/10.1037/rep0000167
14. Giovannetti, A. M., Solari, A., & Pakenham, K. I. (2021). Effectiveness of a group resilience intervention for people with multiple sclerosis delivered via frontline services. *Disability and Rehabilitation*. https://doi.org/10.1080/09638288.2021.1960441
15. Giovannetti, A. M., Messmer Uccelli, M., Solari, A., & Pakenham, K. I. (2022). Evaluation of a program for training psychologists in an acceptance and commitment therapy resilience intervention for people with multiple sclerosis: A single-arm longitudinal design with a nested qualitative study. *Disability and Rehabilitation*. https://doi.org/10.1080/09638288.2022.2025926
16. Lloyd, J., Bond, F. W., & Flaxman, P. E. (2013). The value of psychological flexibility: Examining psychological mechanisms underpinning a cognitive behavioural therapy intervention for burnout. *Work & Stress, 27*(2), 181–199. https://doi.org/10.1080/02678373.2013.782157
17. Puolakanaho, A., Tolvanenb, A., Kinnunena, S. M., & Lappalainena, R. (2020). A psychological flexibility -based intervention for burnout: A randomized controlled trial. *Journal of Contextual Behavioral Science, 15*, 52–67.
18. de Wit, M., & Hajos, T. (2020). Quality of Life. In M. D. Gellman (Ed.), *Encyclopedia of behavioral medicine*. Springer.
19. The World Health Organization Quality of Life assessment (WHOQOL): position paper from the World Health Organization. (1995). *Social Science & Medicine (1982), 41*(10), 1403–1409. https://doi.org/10.1016/0277-9536(95)00112-k
20. Mayo, N. E. (2015). *ISOQOL dictionary of quality of life and health outcomes measurement. Milwaukee: International society for quality of life research* (p. 115). Eds. International Society for Quality of Life Research (ISOQOL).
21. Densham, S., Williams, D., Johnson, A., & Turner-Cobb, J. M. (2016). Enhanced psychological flexibility and improved quality of life in chronic fatigue syndrome/myalgic encephalomyelitis. *Journal of Psychosomatic Research, 88*, 42–47. https://doi.org/10.1016/j.jpsychores.2016.07.009
22. Gloster, A. T., Klotsche, J., Chaker, S., Hummel, K. V., & Hoyer, J. (2011). Assessing psychological flexibility: What does it add above and beyond existing constructs? *Psychological Assessment, 23*(4), 970–982. https://doi.org/10.1037/a0024135
23. Kashdan, T. B., & Rottenberg, J. (2010). Psychological flexibility as a fundamental aspect of health. *Clinical Psychology Review, 30*(7), 865–878. https://doi.org/10.1016/j.cpr.2010.03.001
24. Pakenham, K. I., Tossani, E., Grandi, S., & Landi, G. (2023). The mediating role of psychological flexibility in the relationship between resilience and distress and quality of life in people with multiple sclerosis. *Journal of Health Psychology*. https://doi.org/10.1177/13591053231182364
25. Giovannetti, A. M., Quintas, R., Tramacere, I., Giordano, A., Confalonieri, P., Messmer Uccelli, M., Solari, A., & Pakenham, K. I. (2020). A resilience group training program for people with multiple sclerosis: Results of a pilot single-blind randomized controlled trial and nested qualitative study. *PLoS ONE, 15*(4), e0231380. https://doi.org/10.1371/journal.pone.0231380
26. Francis, A. W., Dawson, D. L., & Golijani-Moghaddam, N. (2016). The development and validation of the Comprehensive assessment of Acceptance and Commitment Therapy processes (CompACT). *Journal of Contextual Behavioral Science, 5*(3), 134–145. https://doi.org/10.1016/j.jcbs.2016.05.003
27. Compston, A., & Coles, A. (2002). Multiple sclerosis. *Lancet (London, England), 359*(9313), 1221–1231. https://doi.org/10.1016/S0140-6736(02)08220-X
28. Aikens, J. E., Fischer, J. S., Namey, M., & Rudick, R. A. (1997). A replicated prospective investigation of life stress, coping, and depressive symptoms in multiple sclerosis. *Journal of Behavioral Medicine, 20*(5), 433–445. https://doi.org/10.1023/a:1025547431847
29. Dennison, L., Moss-Morris, R., & Chalder, T. (2009). A review of psychological correlates of adjustment in patients with multiple sclerosis. *Clinical Psychology Review, 29*(2), 141–153. https://doi.org/10.1016/j.cpr.2008.12.001
30. McCabe, M. P., & McKern, S. (2002). Quality of life and multiple sclerosis: Comparison between people with multiple sclerosis and people from the general population. *Journal of Clinical Psychology in Medical Settings, 9*, 287–295.

31. Meek, C., & das Nair, R., Evangelou, N., Middleton, R., Tuite-Dalton, K., & Moghaddam, N. (2022). Psychological flexibility, distress, and quality of life in secondary progressive multiple sclerosis: A cross-sectional study. *Multiple Sclerosis and Related Disorders, 67*, 104154. https://doi.org/10.1016/j.msard.2022.104154
32. Pakenham, K. I., & Landi, G. (2023). A pilot randomised control trial of an online acceptance and commitment therapy (ACT) resilience training program for people with multiple sclerosis. *Sclerosis, 1*, 27–50. https://doi.org/10.3390/sclerosis1010005
33. Bond, F. W., Hayes, S. C., Baer, R. A., Carpenter, K. M., Guenole, N., Orcutt, H. K., Waltz, T., & Zettle, R. D. (2011). Preliminary psychometric properties of the acceptance and action questionnaire-II: A revised measure of psychological inflexibility and experiential avoidance. *Behavior Therapy, 42*(4), 676–688. https://doi.org/10.1016/j.beth.2011.03.007
34. Doorley, J. D., Goodman, F. R., Kelso, K. C., & Kashdan, T. B. (2020). Psychological flexibility: What we know, what we do not know, and what we think we know. *Social and Personality Psychology Compass*. https://doi.org/10.1111/spc3.12566
35. Kashdan, T. B., Disabato, D. J., Goodman, F. R., Doorley, J. D., & McKnight, P. E. (2020). Understanding psychological flexibility: A multimethod exploration of pursuing valued goals despite the presence of distress. *Psychological Assessment, 32*(9), 829–850. https://doi.org/10.1037/pas0000834
36. Wolgast, M. (2014). What does the acceptance and action questionnaire (AAQ-II) really measure? *Behavior Therapy, 45*(6), 831–839. https://doi.org/10.1016/j.beth.2014.07.002
37. Gámez, W., Chmielewski, M., Kotov, R., Ruggero, C., & Watson, D. (2011). Development of a measure of experiential avoidance: The multidimensional experiential avoidance questionnaire. *Psychological Assessment, 23*(3), 692–713. https://doi.org/10.1037/a0023242
38. Tynan, M., Afari, N., Dochat, C., Gasperi, M., Roesch, S., & Herbert, M. S. (2022). Confirmatory factor analysis of the comprehensive assessment of acceptance and commitment therapy (CompACT) in active-duty military personnel. *Journal of Contextual Behavioral Science, 25*, 115–121. https://doi.org/10.1016/j.jcbs.2022.07.001
39. Hsu, T., Hoffman, L., & Thomas, E. B. K. (2023). Confirmatory measurement modeling and longitudinal invariance of the CompACT-15: A short-form assessment of psychological flexibility. *Psychological Assessment, 35*(5), 430–442. https://doi.org/10.1037/pas0001214
40. Golijani-Moghaddam, N., Morris, J. L., Bayliss, K., & Dawson, D. L. (2023). The CompACT-10: Development and validation of a comprehensive assessment of acceptance and commitment therapy processes short-form in representative UK samples. *Journal of Contextual Behavioral Science, 2023*(29), 59–66. https://doi.org/10.1016/j.jcbs.2023.06.003
41. Giovannetti, A. M., Pöttgen, J., Anglada, E., Menéndez, R., Hoyer, J., Giordano, A., Pakenham, K. I., Galán, I., & Solari, A. (2022). Cross-country adaptation of a psychological flexibility measure: The comprehensive assessment of acceptance and commitment therapy processes. *International Journal of Environmental Research and Public Health, 19*(6), 3150. https://doi.org/10.3390/ijerph19063150
42. MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*(1), 84–99. https://doi.org/10.1037/1082-989X.4.1.84
43. Polman, C. H., Reingold, S. C., Banwell, B., Clanet, M., Cohen, J. A., Filippi, M., Fujihara, K., Havrdova, E., Hutchinson, M., Kappos, L., Lublin, F. D., Montalban, X., O'Connor, P., Sandberg-Wollheim, M., Thompson, A. J., Waubant, E., Weinshenker, B., & Wolinsky, J. S. (2011). Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Annals of Neurology, 69*(2), 292–302. https://doi.org/10.1002/ana.22366
44. Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica, 67*, 361–370. https://doi.org/10.1111/j.1600-0447.1983.tb09716.x
45. Connor, K. M., & Davidson, J. R. (2003). Development of a new resilience scale: The Connor-Davidson resilience scale (CD-RISC). *Depression and Anxiety, 18*(2), 76–82. https://doi.org/10.1002/da.10113
46. Hayes, S. C., Strosahl, K., Wilson, K. G., Bissett, R. T., Pistorello, J., Toarmino, D., Polusny, M. A., Dykstra, T. A., Batten, S. V., Bergan, J., Stewart, S. H., Zvolensky, M. J., Eifert, G. H., Bond, F. W., Forsyth, J. P., Karekla, M., & McCurry, S. M. (2004). Measuring experiential avoidance: A preliminary test of a working model. *The Psychological Record, 54*(4), 553–578. https://doi.org/10.1007/BF03395492
47. Brown, K. W., & Ryan, R. M. (2003). The benefits of being present: Mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology, 84*(4), 822–848. https://doi.org/10.1037/0022-3514.84.4.822
48. Gillanders, D. T., Bolderston, H., Bond, F. W., Dempster, M., Flaxman, P. E., Campbell, L., Kerr, S., Tansey, L., Noel, P., Ferenbach, C., Masley, S., Roach, L., Lloyd, J., May, L., Clarke, S., & Remington, B. (2014). The development and initial validation of the cognitive fusion questionnaire. *Behavior Therapy, 45*(1), 83–101. https://doi.org/10.1016/j.beth.2013.09.001
49. Wilson, K. G., Sandoz, E. K., Kitchens, J., & Roberts, M. (2010). The valued living questionnaire: Defining and measuring valued action within a behavioral framework. *The Psychological Record, 60*, 249–272.
50. Vickrey, B. G., Hays, R. D., Harooni, R., Myers, L. W., & Ellison, G. W. (1995). A health-related quality of life measure for multiple sclerosis. *Quality of Life Research, 4*(3), 187–206. https://doi.org/10.1007/BF02260859
51. Kurtzke, J. F. (1983). Rating neurologic impairment in multiple sclerosis: An expanded disability status scale (EDSS). *Neurology, 33*(11), 1444–1452. https://doi.org/10.1212/wnl.33.11.1444
52. Bollen, K. A. (1989). *Structural equations with latent variables*. John Wiley & Sons.
53. Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research, 18*(3–4), 117–144. https://doi.org/10.1080/03610739208253916
54. 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.
55. Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55.
56. Hu, L., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling* (pp. 76–99). Sage.
57. Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*(3), 464–504.
58. Cohen, J. (1977). *Statistical power analysis for the behavioral sciences. New York: Academic Press Nunnally, J. C. & Bernstein, I. H. (1994) Psychometric Theory* (3rd ed.). McGraw-Hill.
59. Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420–428. https://doi.org/10.1037/0033-2909.86.2.420
60. Steinmetz, H. (2013). Analyzing observed composite differences across groups: Is partial measurement invariance enough? *Methodology, 9*(1), 1–12. https://doi.org/10.1027/1614-2241/a000049
61. Schmitt, N., Golubovich, J., & Leong, F. T. (2011). Impact of measurement invariance on construct correlations, mean differences, and relations with external correlates: An illustrative

example using Big Five and RIASEC measures. *Assessment, 18*(4), 412–427. https://doi.org/10.1177/1073191110373223

62. Coons, S. J., Gwaltney, C. J., Hays, R. D., Lundy, J. J., Sloan, J. A., Revicki, D. A., Lenderking, W. R., Cella, D., Basch, E., & ISPOR ePRO Task Force,. (2009). Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO good research practices task force report. *Value in Health, 12*(4), 419–429. https://doi.org/10.1111/j.1524-4733.2008.00470.x
63. Muehlhausen, W., Doll, H., Quadri, N., Fordham, B., O'Donohoe, P., Dogar, N., & Wild, D. J. (2015). Equivalence of electronic and paper administration of patient-reported outcome measures: A systematic review and meta-analysis of studies conducted between 2007 and 2013. *Health and Quality of Life Outcomes, 13*, 167. https://doi.org/10.1186/s12955-015-0362-x
64. Eremenco, S., Coons, S. J., Paty, J., Coyne, K., Bennett, A. V., McEntegart, D., & ISPOR PRO Mixed Modes Task Force,. (2014). PRO data collection in clinical trials using mixed modes: Report of the ISPOR PRO mixed modes good research practices task force. *Value in Health, 17*(5), 501–516. https://doi.org/10.1016/j.jval.2014.06.005
65. Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology, 63*(7), 737–745. https://doi.org/10.1016/j.jclinepi.2010.02.00

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.