



Published in final edited form as:

J Contextual Behav Sci. 2022 July ; 25: 115–121. doi:10.1016/j.jcbs.2022.07.001.

Confirmatory Factor Analysis of the Comprehensive Assessment of Acceptance and Commitment Therapy (CompACT) in Active-Duty Military Personnel

Mara Tynan^a, Niloofar Afari^{b,c,d}, Cara Dochat^a, Marianna Gasperi^{b,c,d}, Scott Roesch^e, Matthew S. Herbert^{b,c,d}

^aSan Diego State University/University of California, San Diego Joint Doctoral Program in Clinical Psychology, 6363 Alvarado Court, Suite 103, San Diego, CA, 92120, USA

^bVA San Diego Healthcare System, 3350 La Jolla Village Drive., San Diego, CA 92161, USA

^cUniversity of California San Diego, Department of Psychiatry, 9500 Gilman Drive, La Jolla, CA, 92093, USA

^dVA San Diego Center of Excellence for Stress and Mental Health, 3350 La Jolla Village Drive., San Diego, CA 92161, USA

^eDepartment of Psychology, San Diego State University, San Diego, CA, USA

Abstract

Acceptance and Commitment Therapy (ACT) is a third-wave cognitive-behavioral treatment that targets psychological flexibility (PF), or the ability to persist in behavior consistent with values regardless of unwanted private experiences. The growing use of ACT necessitates an accurate assessment of PF. The Comprehensive Assessment of Acceptance and Commitment Therapy (CompACT) is a three-factor measure of PF (Openness to Experience, Valued Action, and Behavioral Awareness) whose psychometric properties have been examined in limited populations. The current study examined the factor structure and psychometric properties of the CompACT in U.S. military personnel who enrolled in a weight management randomized controlled trial.

Military personnel who either failed or were at risk of failing the Navy's physical fitness assessment or had overweight/obese body mass index (BMI; $N = 178$, $M_{\text{age}} = 29.15$ years; $M_{\text{BMI}} = 33.13 \text{ kg/m}^2$; 61.8% female) completed the CompACT and other questionnaires. Confirmatory factor analysis was used to evaluate the three-factor structure of the original 23-item CompACT (CompACT-23) as well as an 18-item version identified in a Portuguese sample (CompACT-18). Internal consistency and convergent validity with measures of weight-related

Corresponding Author: Matthew S. Herbert, VA San Diego Healthcare System/University of California, San Diego, 3350 La Jolla Village Dr. San Diego, CA 92161, Matthew.Herbert2@va.gov, telephone: 858-642-1411.

Declaration of conflicts of interest: Given her role as an Editorial Board Member, Dr. Afari had no involvement in the peer-review of this article and had no access to information regarding its peer-review.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

experiential avoidance, perceived stress, anxiety, depression, PTSD symptoms, and life satisfaction were examined.

The three-factor structure of the CompACT-23 showed poor fit to the data while the fit of the CompACT-18 was acceptable, as indicated by three descriptive indices ($\chi^2/df = 1.73$, RMSEA = 0.069, SRMR = 0.074). All descriptive fit indices in addition to two comparative fit indices (AIC and BIC) indicated improved model fit over the CompACT-23. The CompACT-18 and its subscales exhibited adequate internal consistency ($\alpha = 0.768$ to 0.861) and convergent validity in expected directions with measures of weight-related experiential avoidance, perceived stress, anxiety, depression, PTSD symptoms, and satisfaction with life.

Results support using the refined, English language CompACT-18 as a three-factor measure of PF in populations such as U.S. military personnel who may benefit from weight management intervention. Future research should explore the content validity of the full measure and the removed items. Lost content could mean the CompACT-23 and the CompACT-18 differentially assess PF. Additional studies should examine psychometric properties in large and more diverse samples to further evaluate the appropriateness of the measure across populations. Longitudinal studies are needed to examine test-retest reliability and sensitivity to change.

Keywords

Comprehensive assessment of acceptance and commitment therapy; ACT; Psychological flexibility; Confirmatory factor analysis; Military personnel; Overweight/obesity; Measurement

Introduction

Acceptance and Commitment Therapy (ACT) is a third-wave transdiagnostic cognitive-behavioral approach with the ultimate goal of increasing psychological flexibility (PF), or “the ability to contact the present moment more fully as a conscious human being, and to change or persist in behavior when doing so serves valued ends” (Hayes et al., 2006, p. 7). With the growing use of ACT, it is important to accurately assess PF to verify this purported mechanism of behavior change, inform the refinement of ACT-based therapies, and track treatment progress and outcomes. Further, given the growing literature that ACT may be effective in improving outcomes in a variety of physical and mental health conditions (Gloster et al., 2020; Graham et al., 2016), a reliable and valid general measure of PF can facilitate assessment of this central treatment mechanism.

Increasing PF is accomplished within ACT by targeting six distinct but interrelated treatment processes: acceptance, defusion, self-as-context, present moment awareness, values, and committed action. These processes are often grouped into the “triflex:” (1) openness to experience and detachment from literalness (acceptance; defusion); (2) self-awareness and perspective-taking (present moment awareness; self-as-context); and (3) motivation and activation (values; committed action; (Hayes et al., 2011). A number of existing measures assess PF in specific contexts, such as chronic pain (McCracken & Vowles, 2007) and weight management (Lillis & Hayes, 2008), or attempt to isolate single ACT processes such as experiential acceptance/avoidance (Gámez et al., 2014), cognitive (de)fusion (Gillanders

et al., 2014), and valued action (Wilson et al., 2010). Alternately, the widely-used Acceptance and Action Questionnaire (AAQ-II), aims to assess psychological (in)flexibility more holistically. The AAQ-II was adapted from the original AAQ, which exhibited poor internal consistency and an unstable factor structure (Bond et al., 2011). However, the AAQ-II has been criticized for poor discriminant validity such that it does not distinguish psychological inflexibility from the more general experience of aversive emotions or the distressful response to those emotions (Landi et al., 2021). While psychological inflexibility may be associated with the experience of distress, ACT theory maintains that these are distinct constructs (Bond et al., 2011).

Given these criticisms and the need to assess PF and its subprocesses in broader contexts, there is a strong interest in new and psychometrically valid measures of PF. One of these, the Comprehensive Assessment of Acceptance and Commitment Therapy Processes (CompACT; Francis et al., 2016), assesses PF using the triflex. The CompACT was developed and validated in a non-clinical sample of university students in the United Kingdom ($N = 377$; 74% female, 94.4% White, M_{Age} : 31.34 years). The final 23-item CompACT (CompACT-23) consists of three subscales: openness to experience (OE), behavioral awareness (BA), and valued action (VA). Since its publication, the CompACT-23 has been used in adults with cardiovascular disease and risk (Kibbey et al., 2020), adults in romantic relationships (Rogge et al., 2019), and Finnish university students (Gallego et al., 2020), all of which have shown strong internal consistency for the CompACT-23. However, exploratory factor analysis (EFA) has demonstrated poor structural validity and variation in item difficulty across student, treatment-seeking, and community samples (Ong et al., 2020). Another recent study with the Portuguese translation of the CompACT-23 demonstrated better model fit after excluding five items from the OE subscale, resulting in an 18-item version (Trindade et al., 2021). A test of the three-factor structure with these 18 items reflected a solution explaining a high proportion of the variance (49.81%), with all items loading on the prescribed factors from Francis et al. (2016).

Although the CompACT-23 has gained traction for use in various populations (Gallego et al., 2020; Kibbey et al., 2020; Rogge et al., 2019), it does not yet have a large literature base to support its psychometric validity and subsequent use. The mixed findings indicate the CompACT-23 still needs to be sufficiently examined in more diverse samples to help clarify its utility as a measure of PF. Additionally, different language versions of a measure cannot be assumed to have the same psychometric properties across languages and cultures. Thus, the factor structure of the revised CompACT-18 should be validated using the English language version for use with English-speaking individuals in the U.S. The aims of the current study were to 1) examine the factor structure of both the English language CompACT-23 and CompACT-18 in a sample of U.S. active-duty military personnel enrolled in a randomized controlled trial of an ACT-enhanced weight management intervention, and 2) assess internal consistency and convergent validity of the total score and three subscales with measures of weight-related experiential avoidance, perceived stress, anxiety, depression, PTSD symptoms, and life satisfaction.

Methods

Participants and Procedures

Participants ($N = 178$) were adult active-duty military personnel (98.3% Navy, 1.7% Marine) who failed or were at risk of failing the Navy's physical fitness assessment and/or whose body mass index (BMI) was classified as overweight/obese ($BMI \geq 25 \text{ kg/m}^2$; (Centers for Disease Control and Prevention, 2021). Participants were enrolled in a randomized controlled trial comparing the Navy's standard weight management program ("ShipShape") to an ACT-enhanced version of the program. Data used for this study were collected at baseline. All procedures were approved by the respective Institutional Review Board and Research and Development Committee. All participants provided informed consent.

Measures

Sociodemographic variables—Participants reported age, gender, race/ethnicity, paygrade, and marital status. BMI was calculated using baseline height and weight. Paygrade was assessed according to the individual's rank and corresponding paygrade. Enlisted (E) paygrades range from E1 to E9. Officer (O) paygrades start from O1 and similarly increase with rank.

Psychological flexibility—As aforementioned, the CompACT was used to assess PF. Respondents rated the truth of each statement for themselves from 0 (*strongly disagree*) to 6 (*strongly agree*). As described above, the CompACT contains three subscales: (1) openness to experience (OE; 10 items), (2) behavioral awareness (BA; 5 items), and (3) valued action (VA; 8 items). Example items include, "I can take thoughts and feelings as they come, without attempting to control or avoid them" (OE); "I rush through meaningful activities without being really attentive to them" (BA); "I can identify the things that really matter to me in life and pursue them" (VA). Total and subscale scores are calculated by reverse scoring appropriate items and summing item responses. Higher scores indicate greater PF or greater presence of OE, BA, and VA.

Internal consistency of the CompACT-23 total ($\alpha = 0.91$), and OE ($\alpha = 0.84$ to 0.85), BA ($\alpha = 0.89$ to 0.91), and VA ($\alpha = 0.88$ to 0.90) subscales has been determined as good to excellent in various samples (Rogge et al., 2019) (Ong et al., 2020). The CompACT-23 was administered to participants, of which the 18 items were extracted for the current analyses.

The following constructs were used to assess convergent validity.

Weight-related experiential avoidance—The 10-item Acceptance and Action Questionnaire for Weight-Related Difficulties – Revised (AAQW-R; Palmeira et al., 2016) was used to measure weight-related experiential avoidance (EA). Respondents indicate the truth of each statement on a 7-point scale from 1 (*Never True*) to 7 (*Always True*). Higher scores indicate greater weight-related EA. The AAQW-R total score has demonstrated good internal consistency ($\alpha = 0.73$ to 0.88), stability over time, sensitivity to clinical changes, and measurement invariance across BMI groups (Dochat et al., 2020; Palmeira et al., 2016).

Perceived Stress—The Perceived Stress Scale (PSS-4) is a 4-item survey used to measure the degree to which life situations are considered stressful (Cohen et al., 1983). Respondents indicate how often they felt or thought a certain way over the last month on a 5-point scale from 0 (*Never*) to 4 (*Very Often*). Items are summed to create total scores, with higher scores indicating greater perceived stress. The PSS-4 has demonstrated good internal consistency (Cohen et al., 1983).

Anxiety and Depression Symptoms—The Patient Health Questionnaire for Depression and Anxiety (PHQ-4) is a brief, 4-item screening measure for depression and anxiety (Kroenke et al., 2009). Respondents indicate how often they have been bothered by specific problems on a 5-point scale from 0 (*Not at all*) to 4 (*Nearly every day*). Scores from the first two questions are summed to create an anxiety score (GAD-2) and scores from the second two questions are summed to create a depression score (PHQ-2). Higher scores indicate greater anxiety or depression symptoms. The GAD-2 and PHQ-2 have demonstrated good internal consistency (Kroenke et al., 2009).

Post-Traumatic Stress Symptoms—The Primary Care Screen for PTSD (PC-PTSD) is a brief, 4-item screening measure of post-traumatic stress symptoms (Prins et al., 2016). Respondents indicate if they have had an experience “so frightening, horrible, or upsetting” that they had certain experiences as a result in the past month (*yes/no*). Items are summed to create a total score, with higher scores indicating greater distress from PTSD symptoms. The PC-PTSD has demonstrated good internal consistency (Prins et al., 2016).

Satisfaction with Life—The Satisfaction with Life Scale (SWLS) is a 5-item survey assessing global life satisfaction (Diener et al., 1985). Respondents indicate to what degree they agree or disagree with various statements related to global life satisfaction on a 7-point scale from 1 (*Strongly disagree*) to 7 (*Strongly agree*). Items are summed to create a total score, with higher scores indicating greater satisfaction with life. The SWLS has demonstrated good internal consistency (Diener et al., 1985).

Statistical Analyses

Preliminary—Confirmatory factor analysis (CFA) was conducted using R with the Lavaan package (R Core Team, 2020) and all other analyses were conducted using IBM SPSS Statistics for Macintosh, version 27. Less than 5% of the data were missing and were determined to be missing at random. Mean replacement was used to account for missing data points on measures without specified missing estimation procedures. Item distributions were evaluated for normality and outliers, and multicollinearity among items and subscales was assessed using variance inflation factor ($VIF < |5|$; (Kline, 2005). Multivariate normality was assessed using Mardia’s tests (Mardia, 1970). T-tests and chi-square tests were used to compare sample means across sociodemographic characteristics.

Confirmatory factor analysis—None of the distributions for the observed variables exhibited significant skewness or kurtosis. Based on VIF, multicollinearity was not suspected. The null hypothesis was rejected for Mardia’s tests ($p < .001$), indicating a

violation of multivariate normality (Mardia, 1970). As a result, a robust maximum likelihood estimator (MLR) was used to conduct the CFA.

CFA is a statistical procedure used to assess the quality of the fit of a model or factor structure hypothesized to underly measured variables (Meyers et al., 2006). CFA was used to evaluate the fit of the three-factor structures using maximum likelihood estimation. In addition to the three-factor structure with 23 items proposed by Francis et al. (2016), we tested the three-factor structure of the CompACT-18 with five of the items from the OE subscale removed, as has been suggested by Trindade et al. (2021) in the refinement and validation of the Portuguese version.

To evaluate model fit, the following estimators of goodness of fit were used: χ^2/df (≤ 3 acceptable, ≤ 2 good; Schermelleh-Engel et al., 2003), Comparative Fit Index (CFI > 0.90 acceptable, > 0.95 desirable; (Hu & Bentler, 1998), Tucker-Lewis Index (TLI > 0.90 acceptable, > 0.95 desirable; (Hu & Bentler, 1998), Root Mean Square Error of Approximation (< 0.05 good fit, < 0.08 fair fit, > 0.10 poor fit; (Hu & Bentler, 1999), and Standard Root Mean Square Residual (SRMR < 0.05 good fit, and < 0.08 acceptable fit; (Hu & Bentler, 1999). Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were used to compare models, with lower values indicating better model fit.

Internal consistency and convergent validity—Cronbach's alpha was used to assess internal consistency, with the assumption that $\alpha > 0.80$ would be adequate for research purposes (Lance et al., 2006; Nunally & Bernstein, 1978). Pearson's correlation coefficients were used to assess convergent validity ($r = |0.00-0.10|$ negligible; $r = |0.10-0.39|$ weak; $r = |0.40-0.69|$ moderate; $r = |0.70-0.89|$ strong; $r = |0.90-1.00|$ very strong; (Hinkle et al., 2003). Based on convergent validity analyses carried out by Francis et al. (2016), a moderate to large negative correlation was expected between the CompACT and weight-related EA (AAQW-R) scores, with the OE subscale of the CompACT having the strongest relationship of the subscales. Moderate negative correlations were expected between the CompACT and perceived stress (PSS-4), anxiety (GAD-2), depression (PHQ-2), and post-traumatic stress symptoms (PC-PTSD). Further, the CompACT's OE subscale was expected to have the strongest relationships with each of these measures. Finally, the CompACT total score was expected to have a moderate positive correlation with satisfaction with life (SWLS).

Results

Sociodemographic variables and descriptive statistics

Of the 178 participants (98.3% Navy, 1.7% Marines), more than half were female (61.8%), and the mean age was 29.15 years ($SD = 6.93$). Regarding race and ethnicity, 59.6% of the sample was White, 24.7% Black, 7.3% Asian, 4.5% American Indian, 2.8% Pacific Islander and over a quarter identified as Hispanic (28.7%). The majority were married, partnered, or in a significant relationship (70.8%) and in the enlisted paygrade (93.3%). A total of 15.3% of the sample had a BMI in the overweight range and 81.4% had a BMI in the obese range, and the average BMI was 33.13 kg/m^2 ($SD = 3.89$), which is in the obese range (BMI 30 kg/m^2 ; (Centers for Disease Control and Prevention, 2021). More than 1/3 of the sample (39.9%) reported currently receiving mental health treatment.

Confirmatory factor analyses

Model 1.—The three-factor model of the CompACT-23 was first tested (Table 1). All of the descriptive estimators fell outside of the acceptable range except for χ^2/df ($\chi^2/df = 2.04$, CFI = 0.809, TLI = 0.787, RMSEA = 0.081, SRMR = 0.083). This indicates that the model did not fit the data well descriptively, thus model evaluation did not proceed.

Model 2.—The three-factor model of the CompACT-18 was subsequently tested (Table 1). All of the descriptive fit indices showed improvement from Model 1 and three indicated acceptable or good model fit ($\chi^2/df = 1.73$, CFI = 0.889, TLI = 0.871, RMSEA = 0.069, SRMR = 0.074). Furthermore, Model 2 showed lower AIC and BIC values than Model 1, indicating improvement in the fit of the model, thus model evaluation proceeded.

Standardized factor loadings for all items were large and significant ($ps < 0.001$; Figure 1) for each subscale: OE loadings ranged from 0.551 to 0.867, BA loadings ranged from 0.580 to 0.773, and VA loadings ranged from 0.495 to 0.657. Correlations between factors were weak to moderate: OE and BA ($r = 0.642$), OE and VA ($r = 0.318$), BA and VA ($r = 0.605$), which deviates from prior literature (Francis et al, 2016; Ong et al, 2020).

Internal Consistency and Convergent Validity

Since the CompACT-18 yielded better fit than the CompACT-23, internal consistency and validity results for only the CompACT-18 are presented; however, no notable differences were observed with the CompACT-23. The internal consistency for the CompACT-18 total score ($\alpha = 0.861$) and OE ($\alpha = 0.768$), BA ($\alpha = 0.822$), and VA ($\alpha = 0.802$) subscales was acceptable to good. Correlations of the CompACT-18 total and subscale scores with convergent validity measures are presented in Table 2. Consistent with hypotheses, the CompACT-18 total score was moderately, negatively correlated with AAQW-R, PSS, GAD-2, PHQ-2, and PC-PTSD scores, and moderately positively correlated with the SWLS score. Of the CompACT-18 subscales, the BA and VA subscale scores showed moderate relationships with the AAQW-R, PSS, GAD-2, PHQ-2, and SWLS scores and weak relationships with the PC-PTSD score. The OE subscale showed the weakest relationships with these, except for the PC-PTSD correlation, thus exhibiting the least convergent validity of the subscales. Of the three subscales, BA subscale scores showed the strongest correlations with the AAQW-R, GAD-2, PHQ-2, and SWLS scores.

Discussion

The CompACT was developed to address the need for a reliable and valid measure of PF. This study examined the three-factor structure of both the 23- and 18-item English language versions of the CompACT in a sample of U.S. military personnel enrolled in a weight management intervention. Based on both descriptive and comparative fit indices, the three-factor structure of the CompACT-18 showed improved fit to the data over the three-factor structure of the CompACT-23. Additionally, the CompACT-18 total and subscale scores exhibited acceptable to good internal consistency. Therefore, the present findings suggest this 18-item measure can be used to assess and monitor PF as a function of related

sub-constructs (OE, BA, and VA) in U.S. military personnel who may benefit from weight management intervention.

The CompACT-18 total and subscale scores showed moderate associations in expected directions with weight-related experiential avoidance, perceived stress, anxiety, depression, PTSD symptoms, and life satisfaction, providing support for convergent validity. Importantly, the moderate negative relationship between PF as measured by the CompACT-18 and weight-related experiential avoidance indicates that these are related yet distinct constructs. Contrary to expectations, the OE subscale had the weakest relationships with each of the convergent validity measures compared to BA and VA subscales, except for PTSD symptoms. One reason for this may be the demographic characteristics of the current sample compared to those of the development sample, which would indicate cross-cultural differences in the CompACT-18's measurement of OE. Specifically, compared to the initial, non-clinical sample (Francis et al., 2016), the current sample was less female-skewed with a large portion identifying with a race other than White and over one quarter identifying as Hispanic, with an average BMI in the obese range, and a substantial percentage reporting currently receiving mental health treatment. Another possible reason for this discrepancy is the fact that all five items were removed from the OE subscale, potentially impacting this subscale's convergent validity.

When seeking to psychometrically assess and refine measures, it is necessary to consider the balance between theoretical consistency and statistical stability. Reducing items representing theoretical constructs to numbers and fit indices allows for an objective way of assessing the quality and appropriateness of measures both within and across populations. However, the importance of applying these statistical considerations to theory and existing literature cannot be understated. In the current study, statistically conservative choices were made when conducting CFA and careful consideration was to the implications of these results. While the CompACT-18 showed better statistical fit to the data, consistent with findings from Trindade et al. (2021), there exist theoretical implications for the CompACT's representation of the triflex conceptualization of PF. For example, the exclusion of five items from the original OE subscale must be considered. According to Trindade et al. (2021), the truncated subscale may increase the imbalanced representation of acceptance versus defusion items. During development, Francis et al. (2016) indicated that the OE subscale of the CompACT-23 consisted of eight acceptance and two defusion items. Thus, the CompACT-18 may impact content validity as the openness to experience component of the triflex may now be underrepresented. Of note, all five items (6, 13, 18, 20, and 22) identified for removal by Trindade et al. (2021) also did not perform well in the EFA conducted by Ong et al. (2020) in English-speaking samples. Therefore, it may be helpful for future studies to examine these items and revise or rewrite them. For example, one item removed based on poor factor loading in the Portuguese sample was item 20 "thoughts are just thoughts, they don't control what I do" (Trindade et al., 2021). The poor factor loading of this item in both the current study and in community, student, and treatment-seeking samples (Ong et al., 2020) may indicate that ACT-based language is not clear and easy to understand among certain populations. Alternatively, this item may not reflect acceptance and defusion or OE as intended, perhaps because of wording that could also refer to engaging in valued action (i.e., "control what I do"). Accordingly, in their EFA, Trindade et

al. (2021) found that this item cross-loaded onto both the OE and the VA subscales, as did Ong et al. (2020) in community and treatment-seeking, but not student, samples.

Finally, the CompACT-23 may not perform well statistically due to concerns with the triflex conceptualization of psychological flexibility itself. For example, self-as-context and cognitive defusion both involve detaching from relational responding (Dymond & Roche, 2013). As paired in the triflex, defusion and self-as-context are represented in different subscales (OE and BA, respectively) despite being related constructs. In the CompACT-23 with the full OE subscale, the correlation between the OE and BA subscales was higher than in the CompACT-18 (.8 vs. .6, respectively). Thus, the instability of at least some of the removed items may be due to their relatedness to both defusion in the OE subscale and self-as-context in the BA subscale. Together, this means that while the triflex model makes sense theoretically and may have clinical utility, statistical stability may be difficult to attain.

As ACT and other third-wave cognitive-behavioral therapies continue to gain popularity and use in clinical settings, our findings indicate that the CompACT-18 may be useful in measuring PF for assessing treatment targets and progress, with the benefit of reducing the time and burden for completion in either research or clinical settings. Administering the CompACT-18 before treatment could lend insight into which of the ACT processes may be client strengths or areas for growth. Similarly, regular administration would allow clinicians to track relative change in ACT processes throughout treatment and to assess treatment effectiveness. However, our study findings may inform clinicians' interpretation of PF measures by highlighting the current limitations of the measurement of this construct and cautioning against the exclusive use of self-report PF measures to track treatment outcomes. Therefore, it is important also to evaluate clients' level of treatment progress using other methods of assessment such as client interviews. We also encourage clinicians to consider that both current and prior findings indicate that the CompACT may not measure aspects of PF similarly across cultures and languages (e.g., Portuguese, United Kingdom, U.S.; Ong, 2020, Trindade, 2021), and there is potential variation between other groups (e.g., gender, college students, weight-loss seeking, active-duty military). Thus, clinicians should be aware of cultural considerations when assessing and discussing PF.

There are two recently developed measures of PF that can be compared theoretically and psychometrically to the CompACT. The 60-item Multidimensional Psychological Flexibility Inventory (MPFI) was developed to measure both psychological flexibility and inflexibility and each of their respective six subprocesses (Rolffs et al., 2016). Initial evidence shows that the MPFI has good structural validity (Seidler et al., 2020) and discriminant validity with distress (Landi et al., 2021). The six-item Psy-Flex was developed to be a brief, clinically relevant, and contextually sensitive measure of PF (Gloster et al., 2021). In its development and initial validation, the Psy-Flex has shown good internal consistency across clinical and non-clinical samples, good convergent and divergent validity, discriminant validity between clinical and non-clinical samples, and incremental validity over measures of overall symptomatology. To the best of our knowledge, however, these psychometric properties have not yet been demonstrated in secondary samples. Various considerations inform selection of one self-report measure of PF over another, such as intended use (clinical vs. research), participant/patient burden (number of items), and measurement of subprocesses (presence

of subscales). Future research should compare and refine different PF measures to ensure availability of reliable and valid measures of PF and provide evidence of their comparative utility.

Strengths and Limitations

To our knowledge, this is the only study to validate the English language version of the CompACT-18 and to explore its factor structure in a sample of U.S. military personnel who may benefit from weight management intervention. Additionally, the present convergent validity findings are the only conducted with the English-language CompACT-18 and expand on those conducted in the Portuguese-language version to include broader measures of quality of life and mental health functioning.

Nonetheless, this study has several shortcomings. First, the sample size was relatively small ($N = 178$), which may bias parameter estimates. However, this bias tends to be larger for non-normally distributed item responses than for those that are normally distributed (Hau & Marsh, 2004) and previous studies have completed CFA with similar numbers of items and sample sizes (Lue et al., 2015; Tobias et al., 2017). In addition, cutoffs for CFA sample size are inconsistent and the minimum necessary sample size should be determined based on design characteristics, such as the number of indicators per factor and magnitude of the loadings, as opposed to rules of thumb (Gagne & Hancock, 2006). Our sample exceeds the minimum determined by such characteristics. Second, our data were cross-sectional; thus, we were unable to evaluate psychometric properties such as test-retest reliability and sensitivity to change. Third, we examined the CompACT-18 by administering the CompACT-23, then removing the indicated items. While this was identical to the methodology used by Trindade et al. (2021), it is important to consider item-order effects that may occur when examining a subset of items administered in the context of a longer instrument and how this methodology could produce different results than if the subset was administered in isolation (Schwarz, 1999). More research is needed to replicate these findings using the CompACT-18 administered independently.

Future directions

Future studies should consider the content of the five removed items and the impact on both the total scale and OE subscale. Additionally, more research is needed to confirm the present factor structure and other psychometric properties of the measure. Longitudinal studies using the CompACT-18 are needed to establish sensitivity to change. To establish measurement invariance, future studies should explore PF as measured by the CompACT-18 among various sociodemographic groups (e.g., gender, race/ethnicity, SES, active duty/Veteran status) and the CompACT-18 should be administered to large and diverse samples to continue to add to the literature on its structural validity. Clinically, future work should seek increasingly valid and reliable methods of tracking ACT treatment progress and outcomes.

Conclusions

With the growing use of ACT comes the need for a psychometrically sound instrument to measure PF. The current study supports the use of the English language version of the

CompACT-18 as a three-factor measure of PF in English-speaking samples such as U.S. military personnel who may benefit from weight management intervention. Our findings also indicate the CompACT-18 may be appropriate for researchers and clinicians to evaluate PF moving forward.

References

- Bond FW, Hayes SC, Baer RA, Carpenter KM, Guenole N, Orcutt HK, Waltz T, & Zettle RD (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. 10.1016/j.beth.2011.03.007 [PubMed: 22035996]
- Centers for Disease Control and Prevention. (2021). Defining Adult Overweight & Obesity. Retrieved September 10, 2021 from <https://www.cdc.gov/obesity/adult/defining.html>
- Cohen S, Kamarck T, & Mermelstein R. (1983). A Global Measure of Perceived Stress. *Journal of Health and Social Behavior*, 24(4), 385–396. 10.2307/2136404 [PubMed: 6668417]
- Diener E, Emmons RA, Larsen RJ, & Griffin S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessment*, 49(1), 71–75. 10.1207/s15327752jpa4901_13 [PubMed: 16367493]
- Dochat C, Afari N, Wooldridge JS, Herbert MS, Gasperi M, & Lillis J. (2020). Confirmatory factor analysis of the Acceptance and Action Questionnaire for Weight-Related Difficulties-Revised (AAQW-R) in a United States sample of adults with overweight and obesity. *Journal of Contextual Behavioral Science*, 15, 189–196. 10.1016/j.jcbs.2020.01.006 [PubMed: 32257780]
- Dymond S, & Roche B. (2013). *Advances in relational frame theory: Research and application*. New Harbinger Publications, Inc.
- Francis AW, Dawson DL, & 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. 10.1016/j.jcbs.2016.05.003
- Gagne P, & Hancock GR (2006). Measurement Model Quality, Sample Size, and Solution Propriety in Confirmatory Factor Models. *Multivariate Behavioral Research*, 41(1), 65–83. 10.1207/s15327906mbr4101_5 [PubMed: 26788895]
- Gallego A, McHugh L, Villatte M, & Lappalainen R. (2020). Examining the relationship between public speaking anxiety, distress tolerance and psychological flexibility. *Journal of Contextual Behavioral Science*, 16, 128–133. 10.1016/j.jcbs.2020.04.003
- Gámez W, Chmielewski M, Kotov R, Ruggero C, Suzuki N, & Watson D. (2014). The brief experiential avoidance questionnaire: development and initial validation. *Psychological Assessment*, 26(1), 35–45. 10.1037/a0034473 [PubMed: 24059474]
- Gillanders DT, Bolderston H, Bond FW, Dempster M, Flaxman PE, 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. 10.1016/j.beth.2013.09.001 [PubMed: 24411117]
- Gloster AT, Block VJ, Klotsche J, Villanueva J, Rinner MTB, Benoy C, Walter M, Karekla M, & Bader K. (2021). Psy-Flex: A contextually sensitive measure of psychological flexibility. *Journal of Contextual Behavioral Science*, 22, 13–23. 10.1016/j.jcbs.2021.09.001
- Hau KT, & Marsh HW (2004). The use of item parcels in structural equation modelling: non-normal data and small sample sizes. *Br J Math Stat Psychol*, 57(Pt 2), 327–351. 10.1111/j.2044-8317.2004.tb00142.x [PubMed: 15511312]
- Hayes SC., Luoma JB., Bond FW., Masuda A., & Lillis J. (2006). Acceptance and commitment therapy: model, processes and outcomes. *Behavior Research and Therapy*, 44(1), 1–25. 10.1016/j.brat.2005.06.006
- Hayes SC, Villatte M, Levin M, & Hildebrandt M. (2011). Open, Aware, and Active: Contextual Approaches as an Emerging Trend in the Behavioral and Cognitive Therapies. *Annual Review of Clinical Psychology*, 7(1), 141–168. 10.1146/annurevclinpsy-032210-104449
- Hinkle DE, Wiersma W, & Jurs SG (2003). *Applied statistics for the behavioral sciences* (Vol. 663). Houghton Mifflin College Division.

- Hu L. t., & Bentler PM (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, 3(4), 424–453. 10.1037/1082-989X.3.4.424
- Hu L. t., & Bentler PM (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. 10.1080/10705519909540118
- Kibbey MM, DiBello AM, Babu AA, & Farris SG (2020). Validation of the Valuing Questionnaire (VQ) in adults with cardiovascular disease and risk. *Journal of Contextual Behavioral Science*, 17, 144–151. 10.1016/j.jcbs.2020.07.006
- Kline RB (2005). *Principles and Practice of Structural Equation Modeling*, Second Edition. Guilford Publications. <https://books.google.com/books?id=EkMVZUxZrgIC>
- Kroenke K, Spitzer RL, Williams JBW, & Löwe B. (2009, 2009/11/01/). An Ultra-Brief Screening Scale for Anxiety and Depression: The PHQ–4. *Psychosomatics*, 50(6), 613–621. 10.1016/S0033-3182(09)70864-3 [PubMed: 19996233]
- Lance CE, Butts MM, & Michels LC (2006). The Sources of Four Commonly Reported Cutoff Criteria: What Did They Really Say? *Organizational Research Methods*, 9(2), 202–220. 10.1177/1094428105284919
- Landi G, Pakenham KI, Crocetti E, Grandi S, & Tossani E. (2021). The Multidimensional Psychological Flexibility Inventory (MPFI): Discriminant validity of psychological flexibility with distress. *Journal of Contextual Behavioral Science*, 21, 22–29. 10.1016/j.jcbs.2021.05.004
- Lillis J, & Hayes SC (2008). Measuring avoidance and inflexibility in weight related problems. *International Journal of Behavioral Consultation and Therapy*, 4(4), 348–354. 10.1037/h0100865
- Lue YJ., Wu YY., Liu YF., Lin GT., & Lu YM. (2015). Confirmatory Factor Analysis of the Boston Carpal Tunnel Questionnaire. *Journal of Occupational Rehabilitation*, 25(4), 717–724. 10.1007/s10926-015-9579-0 [PubMed: 25894722]
- Mardia KV (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57(3), 519–530. 10.1093/biomet/57.3.519
- McCracken LM, & Vowles KE (2007). Psychological flexibility and traditional pain management strategies in relation to patient functioning with chronic pain: an examination of a revised instrument. *Journal of Pain*, 8(9), 700–707. 10.1016/j.jpain.2007.04.008 [PubMed: 17611162]
- Nunnally JC, & Bernstein IH (1978). *Psychometric theory*. New York: McGraw-Hill.
- Ong CW, Pierce BG, Petersen JM, Barney JL, Fruge JE, Levin ME, & Twohig MP (2020). A psychometric comparison of psychological inflexibility measures: Discriminant validity and item performance. *Journal of Contextual Behavioral Science*, 18, 34–47. 10.1016/j.jcbs.2020.08.007
- Palmeira L, Cunha M, Pinto-Gouveia J, Carvalho S, & Lillis J. (2016). New developments in the assessment of weight-related experiential avoidance (AAQW-Revised). *Journal of Contextual Behavioral Science*, 5(3), 193–200. 10.1016/j.jcbs.2016.06.001
- Prins A, Bovin MJ, Smolenski DJ, Marx BP, Kimerling R, Jenkins-Guarnieri MA, Kaloupek DG, Schnurr PP, Kaiser AP, Leyva YE, & Tiet QQ (2016). The Primary Care PTSD Screen for DSM-5 (PC-PTSD-5): Development and Evaluation Within a Veteran Primary Care Sample. *Journal of General Internal Medicine*, 31(10), 1206–1211. 10.1007/s11606-016-3703-5 [PubMed: 27170304]
- R Core Team. (2020). R: A language and environment for statistical computing. In R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Rogge RD, Daks JS, Dubler BA, & Saint KJ (2019). It’s all about the process: Examining the convergent validity, conceptual coverage, unique predictive validity, and clinical utility of ACT process measures. *Journal of Contextual Behavioral Science*, 14, 90–102. 10.1016/j.jcbs.2019.10.001
- Rolffs J, Rogge R, & Wilson K. (2016). Disentangling Components of Flexibility via the Hexaflex Model: Development and Validation of the Multidimensional Psychological Flexibility Inventory (MPFI). *Assessment*, 25. 10.1177/1073191116645905
- Schwarz N. (1999). Self-reports: How the questions shape the answers. *American Psychologist*, 54(2), 93–105. 10.1037/0003-066X.54.2.93

- Seidler D, Stone B, Clark BE, Koran J, & Drake CE (2020). Evaluating the factor structure of the Multidimensional Psychological Flexibility Inventory: An independent replication and extension. *Journal of Contextual Behavioral Science*, 17, 23–31. 10.1016/j.jcbs.2020.04.007
- Tobias KG, Lehrfeld J, Rosenfeld B, Pessin H, & Breitbart W. (2017). Confirmatory factor analysis of the Beck Depression Inventory-II in patients with advanced cancer: A theory-driven approach. *Palliative Support Care*, 15(6), 704–709. 10.1017/s1478951517000724
- Trindade IA, Ferreira NB, Mendes AL, Ferreira C, Dawson D, & Golijani-Moghadda 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. 10.1016/j.jcbs.2021.05.002
- Wilson KG, Sandoz EK, Kitchens J, & Roberts M. (2010). The Valued Living Questionnaire: Defining and Measuring Valued Action within a Behavioral Framework. *The Psychological Record*, 60(2), 249–272. 10.1007/BF03395706

Highlights

- The CompACT is a three-factor measure of psychological flexibility
- Psychometric properties were examined in a unique sample of active-duty military seeking weight loss
- Factor structure of the original and refined versions were tested using CFA
- Internal reliability and convergent validity were established
- Refined 18-item version of the CompACT can be used as a three-factor measure of psychological flexibility

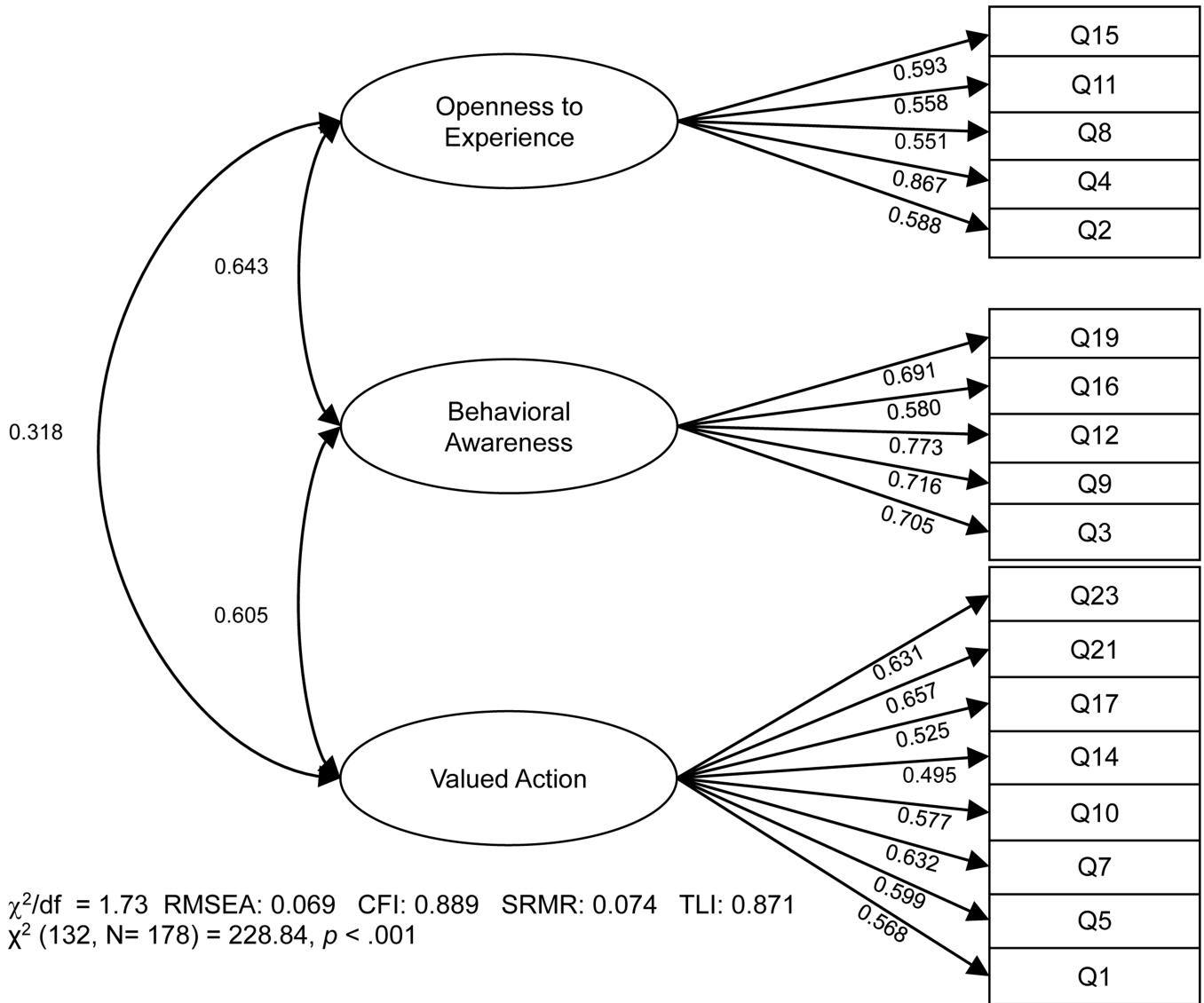


Figure 1. Factor structure and standardized loadings of the 18-item CompACT. Note. RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; SRMR = Standard Root Mean Square Residual; TLI = Tucker Lewis Index; all factor loadings $ps < .001$.

Fit Indices for Tested Models

Table 1.

Model	Chi-square	Df	Chi-square / df	CFI	TLI	RMSEA [90% CI]	SRMR	AIC	BIC
1. 3-factor (23 items)	462.99*	227	2.04	0.809	0.787	0.081 [0.071, 0.092]	0.083	13883.19	14039.10
2. 3-factor (18 items)	228.84*	132	1.73	0.889	0.871	0.069 [0.053, 0.083]	0.074	10819.81	10943.90

Note. Df = degrees of freedom; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation; CI = Confidence Interval; SRMR = Standardized Root Mean-Square Residual; AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion

* $p < 0.001$

Table 2.

Correlations between CompACT total and subscale scores and convergent validity measures

Measure	α	CompACT-18	OE Subscale	BA Subscale	VA Subscale
Weight-related EA AAQW-R	0.886	-0.546	-0.330	-0.527	-0.418
Perceived Stress PSS	0.841	-0.665	-0.412	-0.571	-0.565
Depression/Anxiety	0.903	-0.602	-0.407	-0.528	-0.471
PHQ-2	0.878	-0.571	-0.365	-0.497	-0.469
GAD-2	0.889	-0.550	-0.394	-0.485	-0.407
PTSD Symptoms PC-PTSD	0.822	-0.480	-0.423	-0.389	-0.313
Life Satisfaction SWLS	0.900	0.486	0.243	0.468	0.421

Note. CompACT = Comprehensive Assessment of Acceptance and Commitment Therapy; OE = Openness to Experience; BA = Behavioral Awareness; VA = Valued Action; EA = Experiential Avoidance; AAQW-R = Acceptance and Action Questionnaire for Weight-Related Difficulties-Revised; PSS = Perceived Stress Scale; PHQ = Patient Health Questionnaire; GAD = Generalized Anxiety Disorder; PC-PTSD = Primary Care PTSD Screen; SWLS = Satisfaction with Life Scale; all $ps < .01$.