Abstract

Background Experiential avoidance represents the tendency to avoid negative internal experiences, which is a key concept in Acceptance and Commitment Therapy. However, existing measures of experiential avoidance (i.e., Acceptance and Action Questionnaire II, AAQ-II) have some limitations. This study aims to assess the psychometric properties of the Chinese version of Multidimensional Experiential Avoidance Questionnaire-30 (MEAQ-30) and provide evidence for the reliability and validity of this new instrument.

Methods Two questionnaire surveys were conducted. The first sample (N = 546) was analyzed using classical test theory (CTT), and the second sample (N = 511) was analyzed using multidimensional item response theory (MIRT).

Results CTT supported the six-factor structure of MEAQ-30, indicating good internal consistency and measurement invariance across genders. Furthermore, the Chinese version of MEAQ-30 showed satisfactory convergent and discriminant validity. The incremental validity test showed that after controlling for the effects of neuroticism and AAQ-II, the Chinese version of MEAQ-30 could still significantly predict depression, anxiety, and stress. MIRT indicated that 30 items had good discrimination and difficulty, and the six subscales were sufficiently reliable across the continuum of experiential avoidance.

Conclusion The Chinese version of MEAQ-30 has good reliability and validity and is suitable for assessing experiential avoidance among Chinese populations.

Keywords Experiential avoidance, Classical test theory, Multidimensional item response theory, Measurement invariance
**Introduction**

Experiential avoidance (EA) is defined as the intentional efforts to evade distressing emotions, thoughts, and other private experiences [1, 2]. Although avoiding negative experiences is a natural psychological phenomenon in humans, excessive avoidance of distressing experiences can be counterproductive, which in turn can lead to a range of mental health problems [1, 3]. Numerous studies have demonstrated that high levels of EA are associated with various diagnostic categories, such as anxiety disorder [4], depression disorder [5], and posttraumatic stress disorder [6]. In recent years, there has been a growing interest in studying EA among Chinese populations [7–9]. However, the existing Chinese EA measures are relatively single and vaguely conceptualized, failing to cover the multiple constructs of EA adequately and limiting the validity of EA measures. Therefore, this study aimed to assess the psychometric properties of a new instrument for measuring EA and to provide a more valid measurement for EA-related research in China.

Multiple self-report measures have been created to assess EA. The Acceptance and Action Questionnaire (AAQ-I) is a single-factor instrument initially designed to measure EA, yet it exhibited relatively low internal consistency and retest reliability [10]. Subsequently, Bond et al. [11] revised the AAQ-I to develop the Acceptance and Action Questionnaire (AAQ-II) to address these shortcomings. Like the AAQ-I, the AAQ-II is a single-factor scale. It consists of seven items, each scored on a seven-point scale, with commendable reliability and validity [11]. So far, AAQ-II has found widespread use in Acceptance and Commitment Therapy (ACT) research and has been translated into multiple languages for diverse populations [3, 12, 13]. However, AAQ-II also has some limitations that cannot be ignored. Several studies have indicated that the AAQ-II exhibits low discriminant validity concerning neuroticism, negative affect, and psychological distress [14–16]. Moreover, the unidimensional nature of the AAQ-II fails to capture the full spectrum of EA [3], which may obscure variations in different avoidance tendencies, thereby potentially overlooking valuable and predictive information [17].

Given the broad scope of behaviors, thoughts, and emotions encompassed by EA [1], Gámez et al. [18] developed the Multidimensional Experiential Avoidance Questionnaire (MEAQ) to facilitate the multidimensional assessment of EA. Specifically, the MEAQ comprises 62 items distributed across six dimensions: behavioral avoidance, distress aversion, procrastination, distraction/suppression, repression/denial, and distress endurance. Each dimension represents a distinct behavioral aspect of avoiding distressing experiences. Previous research has consistently demonstrated that MEAQ showed better discriminant validity and reliability than the AAQ-II [18–20]. The multidimensional assessment of EA can help identify potential associations between different facets of EA and various types of psychopathology [21]. However, the MEAQ’s extensive item count limits its practical application. Psychotherapy research requires concise measures to facilitate the repeated assessment of constructs [22].

To simplify the multidimensional assessment of EA, Sahdra et al. [20] streamlined the MEAQ items using genetic algorithms, resulting in the development of the Multidimensional Experiential Avoidance Questionnaire-30 (MEAQ-30). Genetic algorithms are machine learning approaches that mimic Darwinian evolution principles to abbreviate scale in a fully automated manner and maximize the captures of the variance in the original data [20, 23]. MEAQ-30 retained the six dimensions of the original questionnaire and has nearly identical psychometric properties to MEAQ [20], significantly reducing the cognitive burden on participants. To date, MEAQ-30 has been widely used in EA-related fields [19, 24, 25] and translated from English to Turkish [26], which has proved to have sufficient reliability and validity [19, 26]. However, MEAQ-30 has yet to undergo validation in the Chinese context.

Most of the existing studies still use the AAQ-II to assess EA among Chinese populations [7–9] and do not adequately consider the limitations of this instrument, which limits the validity of EA measures. Compared to AAQ-II and MEAQ, MEAQ-30 is undoubtedly a more promising tool, which helps to assess the multidimensional structure of EA quickly. Moreover, few studies have examined the applicability of the multidimensional structure of EA in Eastern cultures. A growing body of studies has shown that EA is influenced by culture and values [3, 27, 28]. Given that Chinese and Western cultures have different perspectives on distress, this may contribute to differences in the expression of EA [3]. Therefore, validating the psychometric properties of the Chinese version of MEAQ-30 will contribute to a clearer understanding of the potential impact of cultural differences on EA and provide a new instrument for advancing research on EA in China.

The current study aims to employ Classical Test Theory (CTT) and Multidimensional Item Response Theory (MIRT) to analyze the psychometric properties of the Chinese version of MEAQ-30. Previous studies have utilized CTT to evaluate the reliability and validity of EA instruments [21, 27, 28]. However, CTT relies on sample-specific estimates for item parameters and reliability, which may not accurately reflect the true level of the trait. On the other hand, Item Response Theory (IRT) models can assess the relationship between item response and latent trait, offering insights into the different levels of the latent trait for each item [29]. Traditional IRT models
are designed for single latent trait instruments, posing challenges in applying unidimensional IRT models to multidimensional data. Multidimensional Item Response Theory (MIRT) models were introduced to overcome this limitation, allowing the simultaneous measurement of multiple traits \[30\]. The multidimensional structure of MEAQ-30 lends itself well to MIRT analysis, enabling the concurrent consideration of correlations between different subscales. To the best of our knowledge, no studies have been conducted to analyze MEAQ-30’s item performance through MIRT.

Taken together, this study aimed to comprehensively assess the psychometric properties of the Chinese version of the MEAQ-30 using both CTT and MIRT. First, the factor structure and internal consistency of the Chinese version of MEAQ-30 were examined. In addition, previous studies did not perform measurement invariance analysis of the MEAQ-30. The current study tested measurement invariance of the MEAQ-30 across genders, examining whether it functions equivalently in males and females. Second, testing the convergent validity and discriminant validity of the MEAQ-30. Considering that existing EA measures (e.g., AAQ-II) overlap with other related concepts (e.g., neuroticism), resulting in poor discriminant validity. Therefore, the current study examined the incremental predictive power of the MEAQ-30 for psychological distress (e.g., depression, anxiety, and stress) after controlling for the effects of neuroticism and AAQ-II. Third, the MIRT was used to analyze item performance (e.g., discrimination and difficulty) in the MEAQ-30 to determine how each item provides information at different levels of EA.

Methods

Participants

The current study conducted two online surveys via the Internet. All participants completed an informed consent form, and the responses were anonymous.

The first survey was conducted between December 2022 and March 2023 and included the Chinese version of MEAQ-30 and other criteria variables (e.g., AAQ-II, Neuroticism, DASS-21). There were 546 participants in this sample from China, ages 18 to 32 ($M=21.36$, $SD=1.99$). Among the participants, 44.14% were male, 55.86% were female, 84.98% were college students, and 15.02% were professional employees.

The second survey was conducted between April 2023 and June 2023 and included the Chinese version of MEAQ-30. There were 511 participants in this sample from China, ages 17 to 38 ($M=22.45$, $SD=3.59$). Among the participants, 43.60% were male, 56.40% were female, 86.70% were college students, and 13.30% were professional employees.

Procedures

The primary author of this study contacted the original author of the MEAQ-30 \[20\] and secured permission to adapt the questionnaire. The forward-backwards translation method was employed to translate the MEAQ-30 into Chinese. Two independent translators translated the MEAQ-30 into Chinese, and after discussions and consensus, an initial draft of the Chinese version of the MEAQ-30 was established. Subsequently, two native English speakers were engaged to translate the Chinese version of the MEAQ-30 back into English. The differences between the original and back-translated versions were discussed in depth until all translators and researchers reached a consensus. Additional file 1 provided the item content for MEAQ-30 (Appendix A is the English version of MEAQ-30, and Appendix B is the Chinese version of MEAQ-30).

Measure

Multidimensional experiential avoidance Questionnaire-30 (MEAQ-30) \[20\]

The MEAQ-30 comprises six subscales: behavioral avoidance, distress aversion, procrastination, distraction/suppression, repression/denial, and distress endurance. Response options for its 30 items range from 1 (strongly disagree) to 6 (strongly agree). Higher scores indicate higher levels of avoidance tendency in the respective subscale. In the current study, Cronbach’s $\alpha$ of the MEAQ-30 was 0.89.

Acceptance and action questionnaire-II (AAQ-II) \[11\]

The AAQ-II is a 7-item scale to measure experiential avoidance/psychological inflexibility (e.g., “I worry about not being able to control my worries and feelings”). Participants responded on a scale of 1 (never true) to 7 (always true). The higher the questionnaire score, the higher the level of experiential avoidance/psychological inflexibility. The Chinese version of the AAQ-II has previously demonstrated satisfactory internal consistency \[31\]. In the present study, Cronbach’s $\alpha$ of the AAQ-II was 0.93.

Chinese big five personality inventory brief version (CBBFPIB) \[32\]: neuroticism scale

The Neuroticism scale within the CBBFPIB comprises eight items (e.g., “I often worry about things that don’t matter”). Participants responded on a scale ranging from 1 (disagree strongly) to 6 (agree strongly). The higher the questionnaire score, the higher the level of neuroticism. In the current study, Cronbach’s $\alpha$ of the Neuroticism scale was 0.91.
**Depression anxiety stress Scale-21 (DASS-21) [33]**

The DASS-21 was utilized to assess psychological distress (e.g., “I felt close to panic”). Responses were gathered on a scale ranging from 0 (did not apply to me at all) to 3 (applied to me very much or most of the time), where higher scores indicate elevated levels of psychological distress. The Chinese version of the DASS-21 has demonstrated satisfactory internal consistency [34]. In the present study, the Cronbach’s α of the DASS-21 was 0.96.

**Data analysis**

CTT analysis was performed on the first sample (N=546) using JASP software (version 0.16.4). First, the mean, standard deviation, skewness, kurtosis, and item-total correlation of the MEAQ-30 items were calculated. Second, the factor structure of the MEAQ-30 was evaluated using confirmatory factor analysis (CFA) and maximum likelihood. Model fit was assessed using the χ²/df ratio, the Comparative Fit Index (CFI), the Tucker-Lewis Index (TLI), the Root-Mean-Squared Error of Approximation (RMSEA), and the Standardized Root Mean Square Residual (SRMR). Acceptable fit criteria include a χ²/ df ratio ≤3, CFI and TLI close to or greater than 0.90, and RMSEA and SRMR lower than or equal to 0.08 [35]. Third, the internal consistency of the MEAQ-30 was assessed using Cronbach’s α coefficient. Values exceeding 0.70 for α indicate good reliability. Fourth, measurement invariance across gender was tested through multigroup CFA analysis, examining configural, metric, and scalar invariance models. Established criteria for invariance included ΔCFI <0.004 and ΔRMSEA <0.050 [36]. Fifth, the convergent validity and discriminant validity of MEAQ-30 are tested using the Pearson correlation, and the incremental validity of MEAQ-30 is tested using hierarchical regression.

MIRT analysis was performed on the second sample (N=511) in this study. The MIRT was implemented using the R package “mirt” [37]. Given the multidimensional structure of the MEAQ-30, the Multidimensional Graded Response Model (M-GRM) [38] was employed. The analysis proceeded as follows: First, the M-GRM’s goodness of fit was assessed using the M2* statistic [39] and its associated RMSEA value. Additional fit indices, such as CFI and TLI, were also obtained. Second, the M-GRM was employed to estimate the discrimination parameter and difficulty parameters of each item. Items with discrimination parameters below 0.24 and above 3.00 may be problematic [40]. Each item included five difficulty parameters (b1, b2, b3, b4, b5) in line with the MEAQ-30 6-point Likert scale. Positive (b) values implied more difficulty, negative (b) values indicated less difficulty, and values close to zero indicated moderate difficulty [41]. Third, the item characteristics curves (ICC), item information curves (IIC), and test information curves (TIC) were presented for a comprehensive understanding.

**Results**

**Descriptive statistics and item-total correlations**

Table 1 presents the descriptive statistics and item-total correlations for the MEAQ-30 items. The skewness values for the 30 items ranged from −1.19 to 0.90, and kurtosis ranged from −1.16 to 2.24. With an absolute value of skewness less than 3 and an absolute value of kurtosis less than 10, it can be concluded that the data is generally acceptable as a normal distribution [42]. The item-total correlations ranged from 0.33 to 0.77, greater than the recommended cut point of 0.30 [43].

**Confirmatory factor analysis**

The fit values for the six-factor measurement model with the MEAQ-30 indicated good fit ($\chi^2$ (390) = 893.644, $\chi^2$/df = 2.291, CFI = 0.929, TLI = 0.921, RMSEA = 0.049, 90% CI [0.044, 0.053], SRMR = 0.057). These findings demonstrated that the MEAQ-30 exhibits good structural validity. Except for item 15, the standardized factor loadings of the other 29 items are all greater than 0.40 (0.50 to 0.84). Item 15 (“I try to deal with problems right away”) belongs to the procrastination subscale and had a relatively low standardized loading (0.34). It should be noted that the content of item 15 is related to procrastination, but its reverse wording could easily lead to acquiescence bias among participants [44]. Moreover, previous studies have shown that adding appropriate reversed items would facilitate a comprehensive assessment of the underlying construct to be measured [45]. Considering that the standardized loading of item 15 was still greater than 0.3, it was decided not to delete this item to ensure consistency with the original questionnaire. The six-factor structure diagram of the Chinese version of MEAQ-30 can be found in Additional File 2.

**Measurement invariance across genders**

The results for measurement invariance are presented in Table 2. The six-factor structure of the MEAQ-30 fitted well in males (N=241) and females (N=305). For gender, the configural model yielded an acceptable fit (CFI = 0.915; RMSEA = 0.054). The metric invariance model demonstrated adequate fit (ΔCFI = 0.000; ΔRMSEA = 0.001). The scalar invariance model also supported measurement invariance (ΔCFI = 0.002; ΔRMSEA = 0.000). Thus, these results suggested that the items of the Chinese version of MEAQ-30 operate nearly identically across genders.
Internal consistency
Cronbach’s α values for the total MEAQ-30 and its subscales ranged from 0.77 to 0.89 (see Table 3). These results suggested that the Chinese version of MEAQ-30 has satisfactory internal consistency.

Convergent and discriminant validity
Pearson correlation analysis was employed to assess the convergent and discriminant validity of the MEAQ-30 (see Table 4). The Chinese version of MEAQ-30 and its six dimensions were significantly positively correlated with the AAQ-II. Except for distraction/suppression, MEAQ-30 and its five dimensions were significantly positively correlated with neuroticism, depression, anxiety, and stress. These results indicated that the Chinese version of MEAQ-30 has good convergent and discriminant validity.

Incremental validity
The incremental validity of the MEAQ-30 was tested by hierarchical regression. As shown in Table 5, after controlling for the effects of neuroticism and AAQ-II,
Table 4: Correlations of the Chinese version of MEAQ-30 with other scales (N=546)

<table>
<thead>
<tr>
<th>Variables</th>
<th>MEAQ-30</th>
<th>BehAvd</th>
<th>DisAvr</th>
<th>Procst</th>
<th>DstSup</th>
<th>RepDny</th>
<th>DisEndr</th>
<th>AAQ-II</th>
<th>CBF-N</th>
<th>DASS-D</th>
<th>DASS-A</th>
<th>DASS-S</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 MEAQ-30</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 BehAvd</td>
<td></td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 DisAvr</td>
<td></td>
<td></td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Procst</td>
<td></td>
<td></td>
<td></td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 DstSup</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 RepDny</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 DisEndr</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 AAQ-II</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 CBF-N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 DASS-D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 DASS-A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>***</td>
<td></td>
</tr>
<tr>
<td>12 DASS-S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>***</td>
</tr>
</tbody>
</table>

Note: BehAvd = behavioral avoidance; DisAvr = distress aversion; Procst = procrastination; DstSup = distraction/suppression; RepDny = repression/denial; DisEndr = distress endurance; CBF-N = neuroticism scale; DASS-D = depression, DASS-A = anxiety, DASS-S = stress; * p < 0.05, ** p < 0.01, *** p < 0.001

M-GRM model fit

MIRT analysis was performed on the second sample (N=511) in this study. At the overall model level, the fitting results of the M-GRM were deemed satisfactory (CFI=0.925, TLI=0.913, RMSEA=0.055, 95% CI [0.050, 0.060]).

Item parameter

Table 6 provided IRT parameter estimates for the six subscales. For behavioral avoidance, the discrimination parameter ranged from 1.63 to 2.30. Distress aversion, the values ranged from 0.95 to 1.81. Procrastination exhibited values ranging from 0.80 to 2.95. Distraction/suppression had values ranging from 1.28 to 1.98. Repression/denial demonstrated values ranging from 1.78 to 2.31. Distress endurance showed values ranging from 1.20 to 2.49. In summary, these findings indicated that all items of the MEAQ-30 exhibited an excellent capacity to discriminate the experiential avoidance dimension. Additionally, all MEAQ-30 items reported difficulty parameters that progressed from less to more difficulty. The difficulty parameters of items were widely distributed, with most items ranging from −3 to 3. Item 15 had a low discrimination parameter, with some difficulty parameters greater than 3, indicating that this item may be problematic.

Item characteristics curves (ICC)

The item characteristics curves are depicted in Fig. 1. This illustration highlights that items with a higher value of discrimination and a steeper curve slope tend to provide information about respondents’ EA within a narrow range [30]. For instance, consider item 14 (“I won’t do something until I absolutely have to”), which exhibits the highest discrimination (a=2.95). In contrast, item 15 (“I try to deal with problems right away”) has the lowest curve slope and the lowest discrimination (a=0.80).

Item information curves (IIC)

The IIC illustrates the amount of information that each item explains concerning the latent trait level [46]. As depicted in Fig. 2, item 14 (“I won’t do something until I absolutely have to”) provided the most information at the medium levels of the latent trait (θ=0). In contrast, item 19 (“When unpleasant memories come to me, I try to put them out of my mind”) offered the most information at lower levels of the latent trait (θ = -2), while item 29 (“I don’t let gloomy thoughts stop me from doing what
I want”) provided the most information at higher levels of the latent trait ($\theta = 2$).

**Test information curves (TIC)**

The test information curves (TIC) were employed to evaluate the discriminative power of six subscales of MEAQ-30 (see Fig. 3). TIC indicated that different subscales provide the most information at varying levels of the latent trait. Specifically, behavioral avoidance (F1) and distraction/suppression (F4) provided the most information at lower latent trait levels and were more appropriate for measuring people with lower levels of EA. Conversely, distress endurance (F6) provided the most information at higher latent trait levels and was more suitable for measuring people with higher levels of EA. Lastly, distress aversion (F2), procrastination (F3), and repression/denial (F5) are more suitable for measuring people with moderate levels of EA.

**Discussion**

This research evaluated the psychometric properties of the Chinese version of MEAQ-30 using both CTT and MIRT analysis. The two methods provided firm evidence for the reliability and validity of the Chinese version of MEAQ-30, indicating that it is a promising instrument for measuring EA among the Chinese population.

The present study demonstrated that the Chinese version of MEAQ-30 exhibits satisfactory structural validity and the same six-factor structure as the other language versions of the MEAQ-30 [20, 26]. Most items had standardized loadings above 0.5, which reflects the EA of the corresponding dimension well. Multiple confirmatory factor analyses supported the measurement invariance of MEAQ-30 across genders. This implied that the items of MEAQ-30 bear consistent meanings for both men and women, allowing for meaningful score comparisons. Furthermore, the internal consistency coefficients for the six dimensions of the Chinese version of MEAQ-30 exceed 0.7, indicating excellent reliability.

The current study showed that the Chinese version of MEAQ-30 had good convergent validity with the AAQ-II, neuroticism, depression, anxiety, and stress. The total MEAQ-30 score and the five dimensions were positively correlated with the above variables. Interestingly, this study found no significant correlation between scores on the distraction/suppression subscale and scores on anxiety and depression. This discrepancy may be interpreted from a cultural perspective. In Chinese culture, the suppression of emotions may be regarded as an adaptive behavior that reduces conflict and promotes social harmony [27, 47, 48]. Consequently, for some Chinese individuals, the distraction/suppression subscale items might be perceived as “not that negative” or “neutral”. They

<table>
<thead>
<tr>
<th>Variables</th>
<th>B</th>
<th>SE</th>
<th>b</th>
<th>T</th>
<th>$R^2$</th>
<th>$\Delta R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>DASS-depression</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step-1</td>
<td>0.26***</td>
<td>0.02</td>
<td>0.56***</td>
<td>13.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neuroticism</td>
<td>0.03**</td>
<td>0.01</td>
<td>0.14**</td>
<td>3.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step-2</td>
<td>0.49***</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neuroticism</td>
<td>0.13***</td>
<td>0.03</td>
<td>0.28***</td>
<td>4.88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAQ-II</td>
<td>0.02</td>
<td>0.01</td>
<td>0.10*</td>
<td>2.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEAQ-30</td>
<td>0.18***</td>
<td>0.03</td>
<td>0.38***</td>
<td>7.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DASS-anxiety</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step-1</td>
<td>0.27***</td>
<td>0.02</td>
<td>0.59***</td>
<td>13.89</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neuroticism</td>
<td>0.03**</td>
<td>0.01</td>
<td>0.11**</td>
<td>2.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step-2</td>
<td>0.50***</td>
<td>0.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neuroticism</td>
<td>0.13***</td>
<td>0.03</td>
<td>0.28***</td>
<td>4.97</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAQ-II</td>
<td>0.02</td>
<td>0.01</td>
<td>0.07</td>
<td>1.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEAQ-30</td>
<td>0.20***</td>
<td>0.03</td>
<td>0.42***</td>
<td>7.86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DASS-stress</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step-1</td>
<td>0.30***</td>
<td>0.02</td>
<td>0.61***</td>
<td>15.23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neuroticism</td>
<td>0.04***</td>
<td>0.01</td>
<td>0.16***</td>
<td>4.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step-2</td>
<td>0.60***</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neuroticism</td>
<td>0.11***</td>
<td>0.03</td>
<td>0.23***</td>
<td>4.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAQ-II</td>
<td>0.03**</td>
<td>0.01</td>
<td>0.11**</td>
<td>2.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEAQ-30</td>
<td>0.26***</td>
<td>0.02</td>
<td>0.50***</td>
<td>10.56</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: *$p<0.05$, **$p<0.01$, ***$p<0.001$
were more likely to agree with the formulations of these items (e.g., “When unpleasant memories come to me, I try to put them out of my mind”), potentially weakening the connection between distraction/suppression and depression and anxiety.

Additionally, the Chinese version of the MEAQ-30 had better discriminant and incremental validity than the AAQ-II. There was a high correlation between AAQ-II and neuroticism ($r = 0.82$), consistent with previous research findings [14, 16], indicating that the concepts measured by AAQ-II and neuroticism overlap. Incremental validity tests showed that after controlling for the effects of neuroticism and AAQ-II on psychological distress, the MEAQ-30 remained an independent predictor variable for psychological distress. It explained a range of additional variance (5–8%) for depression, anxiety, and stress. These findings supported the validity of the Chinese version of MEAQ-30 for measuring EA.

This study proceeded to examine the item parameters of the Chinese version of MEAQ-30 through MIRT. The discrimination parameter of the 30 items ranged from 0.80 to 2.95, indicating that these items had moderate to high discriminatory power [49] and effectively discriminated between participants with different levels of EA. Furthermore, the difficulty parameter for most items ranged from $-3$ to 3, indicating that the items were moderately difficult. Item 15 had the lowest discrimination and the highest difficulty. It provided the least information for procrastination, which is consistent with the results of CTT (e.g., the lowest item-total correlation and standardized factor loading). These results indicated that this item may need to be adjusted.

Item information curves highlighted that item 3, item 10, item 14, item 16, item 22, and item 29 provided higher information in assessing EA dimension. These items can effectively discriminate against individuals with different levels of EA. In contrast, item 2, item 6, item 15, item 18, item 21, and item 26 exhibited lower information in assessing the EA dimension. These six items had relatively low value in evaluating EA, and their item content

<table>
<thead>
<tr>
<th>Item</th>
<th>a1</th>
<th>a2</th>
<th>a3</th>
<th>a4</th>
<th>a5</th>
<th>a6</th>
<th>b1</th>
<th>b2</th>
<th>b3</th>
<th>b4</th>
<th>b5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item1</td>
<td>1.68</td>
<td>0.95</td>
<td>1.80</td>
<td>1.67</td>
<td>1.22</td>
<td>2.32</td>
<td>2.95</td>
<td>0.80</td>
<td>2.17</td>
<td>2.25</td>
<td>2.32</td>
</tr>
<tr>
<td>Item2</td>
<td>1.63</td>
<td>1.80</td>
<td>1.43</td>
<td>1.67</td>
<td>1.43</td>
<td>2.32</td>
<td>2.95</td>
<td>0.80</td>
<td>2.17</td>
<td>2.25</td>
<td>2.32</td>
</tr>
<tr>
<td>Item3</td>
<td>2.30</td>
<td>2.14</td>
<td>2.56</td>
<td>2.78</td>
<td>2.14</td>
<td>2.56</td>
<td>2.78</td>
<td>0.95</td>
<td>2.17</td>
<td>2.25</td>
<td>2.32</td>
</tr>
<tr>
<td>Item4</td>
<td>1.97</td>
<td>2.14</td>
<td>2.56</td>
<td>2.78</td>
<td>2.14</td>
<td>2.56</td>
<td>2.78</td>
<td>0.95</td>
<td>2.17</td>
<td>2.25</td>
<td>2.32</td>
</tr>
<tr>
<td>Item5</td>
<td>1.67</td>
<td>2.14</td>
<td>2.56</td>
<td>2.78</td>
<td>2.14</td>
<td>2.56</td>
<td>2.78</td>
<td>0.95</td>
<td>2.17</td>
<td>2.25</td>
<td>2.32</td>
</tr>
<tr>
<td>Item6</td>
<td>0.95</td>
<td>1.80</td>
<td>1.43</td>
<td>1.67</td>
<td>1.22</td>
<td>2.32</td>
<td>2.95</td>
<td>0.80</td>
<td>2.17</td>
<td>2.25</td>
<td>2.32</td>
</tr>
<tr>
<td>Item7</td>
<td>1.22</td>
<td>1.80</td>
<td>1.43</td>
<td>1.67</td>
<td>1.22</td>
<td>2.32</td>
<td>2.95</td>
<td>0.80</td>
<td>2.17</td>
<td>2.25</td>
<td>2.32</td>
</tr>
<tr>
<td>Item8</td>
<td>1.80</td>
<td>1.43</td>
<td>1.67</td>
<td>1.22</td>
<td>1.80</td>
<td>1.43</td>
<td>1.67</td>
<td>0.95</td>
<td>1.80</td>
<td>1.43</td>
<td>1.67</td>
</tr>
<tr>
<td>Item9</td>
<td>1.43</td>
<td>1.67</td>
<td>1.22</td>
<td>1.80</td>
<td>1.43</td>
<td>1.67</td>
<td>1.22</td>
<td>0.95</td>
<td>1.80</td>
<td>1.43</td>
<td>1.67</td>
</tr>
<tr>
<td>Item10</td>
<td>1.81</td>
<td>2.14</td>
<td>2.56</td>
<td>2.78</td>
<td>2.14</td>
<td>2.56</td>
<td>2.78</td>
<td>0.95</td>
<td>2.17</td>
<td>2.25</td>
<td>2.32</td>
</tr>
</tbody>
</table>

Note: a1 to a6 represent the discrimination parameters of the dimension to which the item belongs; a1 = behavioral avoidance; a2 = distress aversion; a3 = procrastination; a4 = distraction/suppression; a5 = repression/denial; a6 = distress endurance; b1 to b6 represent the difficulty parameter of the items.
may need to be improved. These findings suggested that the Chinese version of MEAQ-30 could be further optimized. Specifically, based on the analysis results of MIRT, the items in the MEAQ-30 could be screened to retain the more informative items, delete the less informative items, and further develop a briefer scale for assessing EA.

Furthermore, the test information curves evidenced the reliability of the six subscales of the MEAQ-30 for measuring the corresponding EA dimensions. It should be noted that the reliability and error of the EA measures vary depending on the dimension. Specifically, behavioral avoidance and distraction/suppression were more appropriate for measuring individuals with lower levels of EA (θ close to -2). As a result, the reliability of estimates is decreased for those with EA levels outside this range. In contrast, distress endurance provided the most information at higher levels of EA (θ close to 2), making this subscale more appropriate for measuring people with higher levels of EA. Additionally, distress aversion, procrastination, and repression/denial provided the most information at the moderate trait level (θ close to 0). These three subscales were more suitable for measuring people with moderate levels of EA. To sum up, these findings contributed to a better understanding of the performance of MEAQ-30 items across different dimensions.

**Implications**

The current study provided solid evidence for the reliability and validity of the Chinese version of MEAQ-30 through CTT and MIRT, making it a valuable tool for assessing EA and opening avenues for further EA-related research in China. It is of significant importance to note that the comprehensive assessment of EA demonstrated by the MEAQ-30 facilitates the identification of potential associations between various avoidance behaviors and other outcomes related to mental health or psychopathology [18, 20]. Furthermore, the current findings highlighted the potential of MIRT to provide additional psychometric insights for MEAQ-30. These findings contributed to a detailed understanding of the item performance of the Chinese version of MEAQ-30.

**Limitations and future studies**

The current study had several limitations. First, the use of self-reporting alone may result in an overestimation of...
Fig. 2  Item information curves for the 30 items

Fig. 3  Test information curves of six subscales of MEAQ-30
the correlation between variables, which could introduce common method biases [50]. Future work should consider validating the MEAQ-30 using diverse sources of information, including experimental designs and clinical ratings. Second, the participant’s sample was predominantly composed of college students. Given the limited number of professional employees in this study, no measurement invariance analysis was conducted between students and employees. Future studies should enhance sample diversity and explore the applicability of MEAQ-30 in other populations. Third, this study lacked a longitudinal assessment. Future longitudinal studies could investigate the causal relationships between EA and various variables.

**Conclusion**

The current study provided firm evidence for the reliability and validity of the Chinese version of MEAQ-30 through CTT and MIRT. Consequently, the Chinese version of MEAQ-30 is a valuable tool for assessing experiential avoidance among Chinese populations.

**Supplementary Information**

The online version contains supplementary material available at https://doi.org/10.1186/s40359-024-01790-x.

**Acknowledgements**

The authors thank all study participants.

**Author contributions**

DX conceived the idea, performed the analysis, and prepared the initial draft of the manuscript. HW, HL, and NW prepared the manuscript for revision. DX, JH, LF, and JS collected the data. XW supervised the findings of this work. All authors discussed the results and contributed to the final manuscript.

**Funding**

This work was supported by the Key Research Base of Humanities and Social Sciences of the Ministry of Education of China (grant number 22JJD190011).

**Data availability**

The data and materials are available from the corresponding author upon reasonable request.

**Declarations**

**Ethics approval and consent to participate**

Approval was obtained from the ethics committee of Tianjin Normal University. The procedures used in this study adhere to the tenets of the Declaration of Helsinki. All the participants completed the informed consent forms before the survey.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare no competing interests.

Received: 9 March 2024 / Accepted: 15 May 2024

**References**


