Journal of Pacific Rim Psychology

www.cambridge.org/prp

The CDA-BPD: retrofitting a traditional borderline personality questionnaire under the cognitive diagnosis model framework

Siwei Peng, Daxun Wang, Xuliang Gao, Yan Cai and Dongbo Tu

School of Psychology, Jiangxi Normal University, Nanchang, China

Original Article

Cite this article: Peng S., Wang D., Gao X., Cai Y., and Tu D. (2019). The CDA-BPD: retrofitting a traditional borderline personality questionnaire under the cognitive diagnosis model framework. *Journal of Pacific Rim Psychology*, Volume 13, e22. https://doi.org/10.1017/prp.2019.14

Received: 16 November 2018
Revised: 2 April 2019
Accepted: 5 April 2019

Keywords:
retrofitting; cognitive diagnosis model; borderline personality disorder; criteria-level information

Author for correspondence: Dongbo Tu, Email: tudongbo@aliyun.com

Abstract

To obtain rich information about the cognitive diagnosis of borderline personality disorder (BPD), this study attempted to retrofit a traditional borderline personality questionnaire so that the improved assessment (called CDA-BPD) could provide more diagnostic information. The retrofitting processes included the following steps: (1) applied an cognitive diagnosis model to analyze the psychometric characteristics of the traditional questionnaire; (2) under the guidance of cognitive diagnosis assessment (CDA), high-quality items were chosen to develop the CDA-BPD and tested on 1,097 subjects; (3) the quality of the CDA-BPD was evaluated; (4) the structure of the CDA-BPD was analyzed. Results indicated that: (1) the CDA-BPD had acceptable reliability and validity; (2) the CDA-BPD had sensitivity of 0.985 and specificity of 0.853 with area under curve (AUC) = 0.956; (3) the two structural factors of the traditional questionnaire were confirmed in the CDA-BPD; χ^2 was 83.01 with $df = 26$, $p < .0001$, comparative fit index (CFI) = 0.97, root mean square error of approximation (RMSEA) = 0.045. It was concluded that the practice of retrofitting a traditional borderline personality assessment for cognitive diagnostic purpose was feasible. Most importantly, under the cognitive diagnosis model framework, CDA-BPD could simultaneously provide general-level information and the detailed symptom criteria-level information about the posterior probability of satisfying each symptom criterion in the *Diagnostic and Statistical Manual of Mental Disorders* (5th edition; *DSM-5*; American Psychiatric Association, 2013) for each individual, which gave further insight into tailoring individual-specific treatments for borderline personality disorder.

Borderline personality disorder (BPD), one of the cluster B Axis II personality disorders in the *Diagnostic and Statistical Manual of Mental Disorders* (5th edition; *DSM-5*; American Psychiatric Association, 2013), is a pervasive pattern of instability of interpersonal relationships, self-image, affects, and marked impulsion, which begins in early adulthood and is considered to be the most prevalent personality disorder in an inpatient setting (Kröger, Huget, & Roepke, 2011). Nine diagnostic criteria of BPD are defined in the *DSM-5*, which specifies the core cognitive, behavioral and interpersonal characteristics for identifying and differentiating BPD from other personality and psychiatric disorders (Aggen, Neale, Røysamb, Reichborn-Kjennerud, & Kendler, 2009). Five or more of these nine criteria must be present for a diagnosis of BPD, according to the *DSM-5*.

To measure and assess BPD, a great number of self-reported instruments have been developed, including the McLean Screening Instrument for Borderline Personality Disorder (MSI-BPD; Zanarini et al., 2003), the Zanarini Rating Scale For Borderline Personality Disorder (ZAN-BPD; Zanarini, 2003), the Borderline Personality Questionnaire (BPQ; Poreh et al., 2006), and the Five Factor Borderline Inventory (FFBI; Mullins-Sweatt et al., 2012). These instruments were established using classical test theory (CTT), in which an individual's test score is determined by recognizing his or her location along a single proficiency continuum (de la Torre & Minchen, 2014), and this score is available for general information about the overall level of symptoms in the spectrum of each instrument taker. Furthermore, these instruments have played an important role in psychological disorder assessments and thus have been popular, partly due to their utility relative to clinical use. However, as some individuals may have different symptoms of BPD although they received the same score, and because the score is only a rough indicator of whether a person has BPD, physicians could not obtain detailed information and were unable to provide targeted treatment for these individuals. Thus, the original intent and design of these assessments could not provide fine-grained diagnostic information and the assessments were not able to provide individual-specific treatments for BPD.

With the developing need for fine-grained information extractable from scale performance data, cognitive diagnosis is now drawing wide attention among researchers and practitioners in psychological disorder measurement, especially due to its potential for diagnosis of individual features and effective individual-specific treatments. Cognitive diagnostic assessments (CDAs; Roberts & Gierl, 2010) aim to provide formative diagnostic feedback through

CAMBRIDGE UNIVERSITY PRESS

fine-grained reporting of an individual's attributes (Jang, 2009). A key aspect of cognitive diagnostic assessments is that cognitive psychology and cognitive diagnosis models (CDMs) are combined within a single framework in these assessments, which enables researchers to assess both general-level diagnostic information and detailed criteria-level information of individuals in a particular assessment domain (Y.W. Lee & Sawaki, 2009). This trend also led to the development of various new psychometric models for cognitive diagnosis that allowed mathematical modeling of individual symptom patterns (de la Torre, 2009; Yi, 2017). Although most of the development and application of CDMs were carried out in the field of education, these models also had sufficient applicability and generality to be applied to the diagnosis of psychological disorders (de la Torre, van der Ark, & Rossi, 2018; Jaeger, Tatsuoka, Berns, & Varadi, 2006; Templin & Henson, 2006).

Nevertheless, it is noteworthy that designing a new cognitive diagnostic assessment for BPD is an expensive, time-consuming and cumbersome process. Because of this, retrofitting provides an effective way to obtain the benefits of CDMs that do not have existing BPD assessments designed directly for cognitive diagnosis purposes. Retrofitting the existing BPD assessments under the CDM framework was a possible approach to improve the original assessment for the cognitive diagnoses to provide more useful information (Chen & Chen, 2016; Liu, Huggins-Manley, & Bulut, 2018; Tu, Gao, Wang, & Cai, 2017).

To investigate how CDM applications could offer specific information for the diagnosis of BPD, the present study aimed to retrofit the traditional BPD questionnaire under the CDM framework. The improved BPD assessment with cognitive diagnostic properties (called CDA-BPD) is the focus of this article: it was able to obtain both general and accurate information for diagnosis and symptom spectrum of BPD, and therapy for BPD. Compared with the conventional self-reported assessments of BPD, the improved assessment in this study was expected to not only estimate the posterior probability of borderline personality disorder (PPBPD) according to the *DSM-5*, but also estimate each individual's unique symptom profile or symptom spectrum, which could potentially increase clinical treatments' effectiveness.

CDM framework for retrofitting

General information for retrofitting the traditional BPD Questionnaire

As mentioned, early self-reported assessments did not provide any further insights into the specific areas of BPD. In order to address this problem, the present study aimed to retrofit the traditional BPD questionnaire under the CDM framework. The improved BPD assessment with cognitive diagnostic properties (CDA-BPD) can measure and evaluate BPD using the symptoms in the *DSM-5*, and offer cognitive diagnosis information beyond what a single test score could provide using these criterion profiles.

The CDA-BPD was initially administered using the traditional BPD instrument – the Borderline Personality Questionnaire (BPQ; Poreh et al., 2006) – an 80-item true/false self-report scale that comprises nine subscales corresponding to the nine BPD criteria of the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text-rev.; *DSM-IV-TR*; American Psychiatric Association, 2000). This study chose the BPQ questionnaire for the following reasons: (1) The BPQ is a representative and widely used questionnaire, and researchers have confirmed that it is a reliable and valid tool to assess borderline personality features (Fonseca-Pedrero, Paino, Lemos-Giráldez, Sierra-Baigrie, & Muñiz, 2011); (2) The BPQ is a true/false self-reported instrument that exactly meets the requirement of the CDM for dichotomic data. Although the BPQ is based on *DSM-IV-TR* criteria, many studies report that there are no differences between the definitions of the nine symptoms of BPD in the *DSM-IV-TR* and *DSM-5* (Conway, Hammen, & Brennan, 2012; Iacono, 2013), so the BPQ was still of significance for the current measurement of BPD.

Additionally, a crucial step to conduct the CDA-BPD was to construct a Q-matrix, which played an important role because it describes what attributes or symptom criteria are measured by each item. Let q_{jk} denote the element in row j and column k of a $J \times K$ Q-matrix, where J and K represented the numbers of items and attributes/symptom criteria respectively. The element q_{jk} was specified to be 1 if the kth attribute was required to answer item j correctly, and zero otherwise. The quality of the Q-matrix determines the quality of the estimated diagnostic model, and a poorly created Q-matrix provides less informative diagnostic or classification indices (Tatsuoka, 1990). In the present study, in order to guarantee the highest quality, the Q-matrix of the CDA-BPD was directly based on the primal BPQ scale (Poreh et al., 2006), which describes the relationship between 80 items and nine attributes/symptom criteria. Appendix A shows the initial Q-matrix of the BPQ scale used in this article. Each row of the Q-matrix pertains to an item, and each column to a symptom. For example, the first row denotes that item 1, "I often do things without thinking them through", measures "Impulsivity in at least two areas that are potentially self-damaging (e.g. spending, sex, substance abuse, reckless driving, binge eating)" (Symptom 4), while item 2, "I often become depressed or anxious out of blue", measures "Affective instability due to a marked reactivity of mood (e.g. intense episodic dysphoria, irritability, or anxiety usually lasting a few hours and only rarely more than a few days)" (Symptom 6). It is critical that all CDMs are equivalent when all items have a simple structure.

Cognitive diagnosis model: G-DINA model

The aim of CDMs is to build close connections between individuals' item responses and their attribute patterns or symptom profiles. In the last few years, several CDMs have been proposed (Templin & Henson, 2010). Those vary in the way that attributes are combined and formalized to estimate the probability of item responses. Most of the applications of CDMs in psychological assessment in particular have involved deterministic inputs, the noisy "and" gate model (DINA; Junker & Sijtsma, 2001), the additive cognitive model (ACDM; de la Torre, 2011), and the reduced reparametrized unified model (RRUM; Hartz, 2002). To synthesize the various CDMs, de la Torre (2011) proposed a general model, the G-DINA (generalized deterministic inputs, noisy "and" gate) model, which allowed reformulating many of the existing CDMs as special cases of this general model.

In the G-DINA model, the required attributes for item could be represented by the reduced vector $\alpha_{lj}^{*} = (\alpha_{l1}, ..., \alpha_{lK_j^*})'$, where $l = 1, ..., 2^{K_j^*}$ and $2^{K_j^*}$ represented the number of unique attribute patterns. The probabilities that examinees with reduced attribute vector α_{lj}^{*} would answer item j correctly was expressed as $P(X_j = 1|\alpha_{lj}^{*})$. The formulation of the G-DINA model in its saturated form (i.e. no restrictions were made) was

https://doi.org/10.1017/prp.2019.14 Published online by Cambridge University Press

$$P(X_j = 1|\alpha_{lj}^*) = \delta_{j0} + \sum_{k=1}^{K_j^*} \delta_{jk}\alpha_{lk} + \sum_{k'=k+1}^{K_j^*} \sum_{k=1}^{K_j^*-1} \delta_{jkk'}\alpha_{lk}\alpha_{lk'} \ldots + \delta_{j12\ldots K_j^*} \prod_{k=1}^{K_j^*} \alpha_{lk}. \quad (1)$$

The function above was the sum of the effects based on the presence of specific attributes and their interactions. Specifically, δ_{j0} represented the baseline probability of a correct response when none of the required attributes was mastered; δ_{jk} was the main effect due to attribute α_k, or the change in the probability of a correct (or most effective) response when a single attribute (i.e. α_k) was mastered; $\delta_{jkk'}$, a first-order interaction effect, represented the change in the probability of a correct response when both α_k and $\alpha_{k'}$ were mastered; and $\delta_{j12\ldots K_j^*}$ was the interaction effect of all required attributes, or the change in the probability of a correct response when all required attributes were needed (de la Torre, 2011). The parameters of the G-DINA model could be estimated using the marginal maximum likelihood estimation (MMLE) algorithm.

The saturated model based on the above function covered all possible item effects (e.g. intercept, main, and interaction effects). By constraining some item effects of the saturated forms, the G-DINA model could be transformed into different reduced CDMs. For example, the DINA model (Junker & Sijtsma, 2001) only included the intercept and final interaction effects, and it was expressed as

$$P(X_j = 1|\alpha_{lj}^*) = \delta_{j0} + \delta_{j12\ldots K_j^*} \prod_{k=1}^{K_j^*} \alpha_{lk}. \quad (2)$$

The ACDM model (de la Torre, 2011) is a special case of the G-DINA model by supposing no interaction effects, and it was formulated as

$$P(X_j = 1|\alpha_{lj}^*) = \delta_{j0} + \prod_{k=1}^{K_j^*} \delta_{jk}\alpha_{lk}. \quad (3)$$

The RRUM model (Hartz et al., 2002) is the log-link G-DINA model without interaction terms, and it was defined as

$$\log\left[P(X_j = 1|\alpha_{lj}^*) = \delta_{j0} + \prod_{k=1}^{K_j^*} \delta_{jk}\alpha_{lk}\right]. \quad (4)$$

More technical details about the saturated and reduced CDMs can be found in de la Torre (2011) and de la Torre and Chen (2011).

Symptom criteria of BPD in the DSM-5

The *DSM-5* defines the main features of BPD, and for an individual to be classified as BPD, he or she must satisfy five or more symptom criteria in Table 1. It should be emphasized that the *DSM-5* diagnostic criteria for BPD has been widely used in academic research and clinical diagnostic practice (Bach, Sellbom, Bo, & Simonsen, 2016; Ferrer et al., 2018; Miller, Morse, Nolf, Stepp, & Pilkonis, 2012).

In this context, with the response data, CDM could estimate two kinds of probabilities for each individual, specifying the posterior probability of satisfying each criterion according to the *DSM-5* (Templin & Henson, 2006), and the posterior probability of meeting five or more criteria of borderline personality disorder (PPBPD), which was also the probability of BPD in the *DSM-5*. Specifically, if an individual had a PPBPD greater than .5, then the individual could be classified as BPD. It should be noted that the posterior probability of satisfying each criterion for each individual showed differences in why an individual received a diagnosis of BPD; for example, two individuals who were both diagnosed as BPD may have diverse pathogeny paths for the disorder because they had different probabilities of symptom criteria. Specifically, except when several other criteria of BPD are met, the first individual may be very sensitive to environmental circumstances, and they may experience intense abandonment fears and inappropriate anger even when faced with a realistic time-limited separation or when there are unavoidable changes in plans (S1). The first individual also may display recurrent suicidal behavior, gestures, threats, or self-mutilating behavior (S5). The second individual may differ from the first individual in that he or she may be troubled by chronic feelings of emptiness (S7) and may frequently express inappropriate, intense anger or have difficulty controlling their anger (S8). Both individuals may be diagnosed as BPD, but the expressions of their disorders are the result of distinct behaviors. The probability of any individual meeting any defined criterion could be estimated by the CDM.

Table 1. Symptom criteria of BPD defined in the *DSM-5*

ID	Symptom criteria
S1	Frantic efforts to avoid real or imagined abandonment. (Note: Do not include suicidal or self-mutilating behavior covered in Criterion 5.)
S2	A pattern of unstable and intense interpersonal relationships characterized by alternating between extremes of idealization and devaluation.
S3	Identity disturbance: markedly and persistently unstable self-image or sense of self.
S4	Impulsivity in at least two areas that are potentially self-damaging (e.g. spending, sex, substance abuse, reckless driving, binge eating). (Note: Do not include suicidal or self-mutilating behavior covered in Criterion 5.)
S5	Recurrent suicidal behavior, gestures, or threats, or self-mutilating behavior.
S6	Affective instability due to a marked reactivity of mood (e.g. intense episodic dysphoria, irritability, or anxiety usually lasting a few hours and only rarely more than a few days).
S7	Chronic feelings of emptiness.
S8	Inappropriate, intense anger or difficulty controlling anger (e.g. frequent displays of temper, constant anger, recurrent physical fights).
S9	Transient, stress-related paranoid ideation or severe dissociative symptoms.

https://doi.org/10.1017/prp.2019.14 Published online by Cambridge University Press

A cognitive diagnosis structure of symptom criteria for the CDA-BPD

The present study not only focused on the evaluation of BPD, but also investigated the potential structure of psychological traits. The underlying structure of the CDA-BPD was based on one explored by Poreh et al. (2006). According to Poreh et al., BPD has two underlying traits: identity/interpersonal and impulsivity. The identity/interpersonal factor includes abandonment (S1), relationships (S2), self-image (S3), affective instability (S6), emptiness (S7), and quasi-psychotic states (S9). The impulsivity factor includes impulsivity (S4), suicide (S5), and intense anger (S8). This study applied a CDM to verify the structure between the two factors and nine symptom criteria. Because the CDM could estimate the presence or absence of a set of criteria, it was possible to model the association between these criteria and verify the structure suggested by Poreh et al. (2006). According to Templin and Henson (2006), a method to model the probability of class membership is the concept of tetrachoric correlations to place the structure on the joint distribution of the attributes. A tetrachoric correlation is the correlation between two underlying normally distributed variables that have both been dichotomized by respective cut-point parameters. This study dichotomized each criterion by using the marginal probability of all participants for each criterion as a cut-point. Therefore, a set of cut-point parameters was estimated. The cut-point parameters were essential because they could be converted to a proportion (or percentage) representing the marginal level of presence for each attribute (or, in the study, satisfaction of each criterion). Such parameters provided information regarding the base rate of each attribute in the population represented by the data set (Templin & Henson, 2006).

Measures

Borderline Personality Questionnaire (BPQ; Poreh et al., 2006) is an 80-item true/false self-report measure that assesses borderline personality traits with nine subscales. Results (Poreh et al., 2006) show that the BPQ has high internal consistency (Kuder-Richardson coefficient = 0.94) and good discriminant validity.

CDA-BPD. The initial item pool for the CDA-BPD contained all 80 items from the BPQ (Poreh et al., 2006). The CDA-BPD items were first translated into Chinese, called the first draft. It was then given to four assessors, including two experts in English and two experts in psychology/psychometrics. Their task was to judge the accuracy of translation and relevance/suitability of each item for measuring BPD in the socio-cultural context of China, and this formed the second draft. Then the study applied 30 Chinese college students to verify the quality and accuracy of the second draft by cognitive interview. Finally, the Chinese version of the CDA-BPD was formed. The study used the G-DINA model to analyze the psychometric characteristics of each item in the CDA-BPD. Furthermore, the study selected high-quality items to develop the final CDA-BPD, which included 55 items. The reliability and validity of the final CDA-BPD will be introduced in detail in the results section of this article.

McLean Screening Instrument for Borderline Personality Disorder (MSI-BPD; Zanarini et al., 2003) is a 10-item, self-report screening measure. According to Zanarini et al. (2003), the MSI-BPD has good test–retest reliability ($r = .72$), high internal consistency ($\alpha = .74$), and acceptable item-total correlation (range .45–.63). A score of 7 or more is the optimal clinical cut-off (Zanarini et al., 2003). Wang, Liang, and Zhong (2008) revised the MSI-BPD in Chinese and the results showed that the Chinese version of MIS-BPD had good reliability and validity. In this study, the Chinese version of MIS-BPD was administered and it had a Cronbach's alpha of .73 and a split-half of .72 in the current study.

Personality Diagnostic Questionnaire-Fourth Edition-BPD Scale (PDQ-4-BPD; Hyler, 1994) contained nine true/false items targeting the BPD criteria. As mentioned by Hyler (1994), a cut-off score of ≥5 indicates the presence of BPD traits. Yang et al. (2000) developed the Chinese version of Personality Diagnostic Questionnaire (PDQ-4) and the results showed that the reliability and validity were acceptable. In this study, the Chinese version of PDQ-4-BPD was applied and it had a Cronbach's alpha of .71 and a split-half of .68 in the current study.

The present study used the Chinese version of the MIS-BPD (Wang et al., 2008) and PDQ-4-BPD (Yang et al., 2000) as an evidence to support the convergent validity of the CDA-BPD for two reasons: (1) they are both self-reported scales of BPD based on the *DSM* diagnostic system and are widely used in Chinese BPD research; (2) the Chinese version of the MIS-BPD (Wang et al., 2008) and the PDQ-4-BPD (Yang et al., 2000) have good reliability and validity that have been demonstrated in many studies (H. Chen, Zhong, & Liu, 2011; Huang et al., 2014).

Participants

The total sample came from 20 Chinese universities and consisted of 1,097 participants who agreed to take part in the study after being informed that their personal information would be kept secret and the test would occupy them for about half an hour. The sample had a mean age of 20.63 years ($SD = 1.524$, range = 18–24) and included 431 males (39.3%) and 666 females (60.7%). The total sample included two parts: one part had 1,029 participants who were used to calibrate the item parameters of the CDA-BPD via CDM; the other part was comprised of 68 participants who were used as a validation sample to assess the sensitivity and specificity of the CDA-BPD.

The validation sample included two paired groups according to the demographic characteristics: BPD group ($N_1 = 34$) and health group ($N_2 = 34$).

BPD group. The BPD group had a score ≥7 for the MIS-BPD and a score ≥5 for the PDQ-4-BPD. Participants also reported that their emotional and behavioral instability seriously affected their daily lives and activities. The BPD group also met the exclusion criteria: (1) ever satisfied the *DSM-5* criteria for schizophrenia or schizoaffective disorder; (2) ever had a history of psychiatric treatment, drinking, or drug abuse; or (3) ever been diagnosed with any brain disease caused by infection, trauma, tumor, inheritance or vascular disease. The BPD group comprised 41.2% males (58.8% females), with 43.2% from city (56.8% from country); 44.1% was a freshman at university (11.8% sophomore, 17.6% junior, 26.5% senior), and had a mean age of 20.62 years ($SD = 1.60$, range = 18–24).

Healthy group. The healthy group had a score of MIS-BPD less than 7 and a score of PDQ-4-BPD less than 5. Moreover, each individual in the healthy group must be paired one by one with the BPD group according to age, sex, grade, and where they live. For example, if there was a participant in the BPD group who was 20 years old, male, a freshman in college and from city, the participant in the healthy group must have the same demographic variables (20 years old, male, a freshman in college and from the city). The healthy group was recruited according to the following exclusion criteria: (1) ever been diagnosed with severe somatic

https://doi.org/10.1017/prp.2019.14 Published online by Cambridge University Press

diseases (e.g. chronic diseases, diabetes, cancer); (2) ever had treatment for psychiatric illness over the past year. Written informed consent was obtained from all participants in this study.

Statistical analysis

The statistical analysis consisted of five steps as follows.

Step 1: Use the G-DINA model to analyze the psychometric characteristics of each item in the CDA-BPD. Analyzing psychometric characteristics is very important in assessing BPD. The psychometric characteristics of each item based on the CDM in this study included item-fit, differential item functioning (DIF), and item discrimination. The study used the $S\text{-}X^2$ item fit statistic (Orlando & Thissen, 2003) to examine item fit, the Wald test statistic (Hou, de la Torre, & Nandakumar, 2014) to detect DIF in different groups (e.g. female and male; rural and urban), and applied the $Disc_j$ index to analyze the item discrimination. For the CDMs, two discrimination indexes were proposed: one was suggested by de la Torre (2008) and the other by de la Torre and Chiu (2016). Each item in the CDA-BPD measured one attribute; therefore, the above two indexes were equivalent. Formula (5) was the index of de la Torre (2008), shown as,

$$Disc_j = P(X_j = 1|\alpha_{lj}^* = 1) - P(X_j = 1|\alpha_{lj}^* = 0), \quad (5)$$

where

$$P(X_j = 1|\alpha_{lj}^* = 1) = 1 - P(X_j = 0|\alpha_{lj}^* = 1) = 1 - s_j, \quad (6)$$

and

$$P(X_j = 1|\alpha_{lj}^* = 0) = g_j. \quad (7)$$

In the above function, s_j (slip parameters) referred to the probability that an individual with all required attributes provided an incorrect response, g_j (guessing parameters) referred to the probability that an individual without all required attributes provided a correct response. $P(X_j = 1|\alpha_{lj}^* = 1)$ referred to the response probability for individuals who present with all the symptom criteria measured by item j, and $P(X_j = 1|\alpha_{lj}^* = 0)$ was the response probability for individuals who did not present any symptom criteria measured by item j. $Disc_j$ was a comprehensive index of slip and guessing parameters. The higher value of $Disc_j$ indicated that the quality of the item was better and could discriminate different individuals.

Step 2: Select high-quality items to develop the final CDA-BPD. The selection of high-quality items was based on the statistical indexes, including discrimination, model-fit in item level, and DIF in Step 1. Any item with low discrimination ($Disc_j < 0.2$), poor item fit ($p < .01$) or having DIF ($p < .01$) was excluded. To obtain the final CDA-BPD with high-quality items, this procedure was repeated until no item was excluded. All the analyses of the CDM were based on the GDINA R package (Version 3.5.1 64-bit; Ma & de la Torre, 2016).

Step 3: Examine the reliability and validity of the final CDA-BPD. To examine the reliability of the CDA-BPD, the study used the coefficients of Cronbach's alpha and Guttman split-half reliability, which were based on the CTT framework; under the framework of the CDM, the symptom-level classification consistency and classification accuracy reliability indices (Cui, Gierl, & Chang, 2012), and Templin's attribute reliability, which was proposed by Templin and Bradshaw (2013), were also evaluated for the CDA-BPD. These reliability indicators could estimate how accurately a CDM classified subjects into correct attribute profiles. Besides, the convergent validity of the CDA-BPD was quantified by the correlation between the CDA-BPD and the Chinese version of the MIS-BPD and PDQ-4-BPD. The content validity of CDA-BPD was also investigated by checking whether the CDA-BPD contained all domains of BPD.

Step 4: Estimate the sensitivity and specificity of the final CDA-BPD. The accuracy of an instrument was commonly assessed by calculating the instrument's sensitivity and specificity. In the present study, the sensitivity of the CDA-BPD was defined as the percentage of those who truly had BPD that was correctly identified by the assessment. It was also sometimes called the true positive rate. The specificity of the CDA-BPD was defined as the percentage of those who truly did not have BPD that was correctly identified by the assessment. The specificity was called the true negative rate. In addition, the study estimated the PPBPD, which was based on the CDA-BPD and *DSM-5* via the CDM, and it meant that individuals who had over 0.5 PPBPD were defined as BPD. Clearly, if an assessment was perfect for its domain, it would be expected to have both the sensitivity and specificity equal to 100%. Moreover, ROC curve analyses were used in the study to evaluate the predictive effectiveness of the CDA-BPD. The perfect assessment had the area under the ROC curve (area under the curve [AUC]) of 1.00, while an assessment with no predictive validity had an AUC of .50 and a linear ROC curve that paralleled the diagonal. Although considered rather arbitrary, AUC values of .60 to .70, .70 to .90, and those above .90 were generally associated with an acceptable or fair, good, and excellent discriminant assessment respectively (Swets, 1988). Sensitivity and specificity evidences of CDA-BPD were collected by using a validation sample in the current study.

Step 5: Analyze the latent structure of symptom criteria for the final CDA-BPD. The estimated structure of the CDA-BPD was based on the model developed by Poreh et al (2006). In their development, the authors used the unidimensional sum scores of items to represent latent factor scores for each criterion. The hypothesized structure mapped the latent criteria onto identity/interpersonal (symptom criterion 1, 2, 3, 6, 7 and 9) and impulsivity (symptom criterion 4, 5, and 8).

In the current study, the method used to analyze the structure of the CDA-BPD was different from Poreh et al (2006). Specifically, the nine criteria were treated as dichotomous variables by nine respective cut-point parameters (the study dichotomized each criterion by using the marginal probability of all participants for each criterion as a cut-point). These cut-point parameters were able to be converted to a proportion representing the marginal level of presence each criterion. Then the tetrachoric correlations among pairs of nine criteria were analyzed and this was used to verify the structure of symptom criteria for the CDA-BPD. And the method mentioned above was referred to Templin and Henson (2006).

The ratio of chi-square to the degrees of freedom (*df*) was calculated in addition to chi-square to evaluate model fitting, with ratios between 1.5 and 5.0 indicating acceptable fit (Byrne, 2012). Additional parameters for fit estimation were: the normal fit index (NFI), the non-normal fit index (NNFI), the comparative fit index (CFI), the root mean square error of approximation (RMSEA), the goodness of fit index (GFI), the adjusted goodness of fit index (AGFI). RMSEA values of 0.08 or lower and NFI, NNFI, CFI, GFI, AGFI values of 0.90 or higher were considered acceptable (Browne

https://doi.org/10.1017/prp.2019.14 Published online by Cambridge University Press

& Cudeck, 1993; Hu & Bentler, 1999). Criterion level confirmatory factor analysis (CFA) was carried out with LISREL 8.0 (Joreskog & Sorbom, 1996) in the current study.

Results

Item analysis of the CDA-BPD

Following the removal of items according to the statistical guidelines (see Step 2 in the statistical analysis section), 55 items were finally included for the CDA-BPD, which are displayed in Table 2. The final 55 items fitted the CDM (i.e. G-DINA model) well ($p > .01$), had no DIF ($p > .01$) on sex group or region group, and had a mean discrimination of 0.37 ($SD = 0.08$). Out of 55 items, 38.2% items had an "excellent" level of discrimination (≥0.40) and 40.0% items had a "good" level of discrimination (0.30~0.39), which indicated that the remaining 55 items of the CDA-BPD had a high discrimination of item response probability between participants who met or did not meet symptom criteria assessed by these items. Furthermore, the final CDA-BPD assessed all nine symptom criteria for BPD defined in the *DSM-5*. The number of items assessed each symptom level ranged from 4 to 10 with an average of 7.

Reliability and validity

The reliability of the CDA-BPD was investigated based on both CTT and CDM frameworks. Under the CTT framework, the Cronbach's alpha coefficient of the CDA-BPD was .914 and the Guttman split-half was .880. However, in the CTT framework, the traditional statistical methods could not estimate the reliability for each symptom criterion. CDM represented a new way for estimating the reliability of these nine symptom criteria in *DSM-5*. Results showed that under the CDM framework, the classification consistency reliability of nine attributes ranged from 0.919 to 0.946 with an average of 0.933, and the classification accuracy reliability of nine attributes ranged from 0.929 to 0.955 with an average of 0.941. Besides, Templin's attribute reliability of nine attributes ranged from 0.925 to 0.980 with an average of 0.964. These results indicate that the reliability of CDA-BPD was acceptable.

As for the convergent validity of the CDA-BPD, the test score of the CDA-BPD was significantly correlated with the scores of the MIS-BPD ($r = .629$, $p < .01$) and PDQ-4 ($r = .664$, $p < .01$). It was also found that the estimated PPBPD of the CDA-BPD was significantly correlated with the scores of the MIS-BPD ($r = .613$, $p < .01$) and PDQ-4 ($r = .637$, $p < .01$).

To further examine the CDA-BPD validity, cross-validation was examined using a validation sample that consisted of a healthy control group ($N_1 = 34$) and a BPD group ($N_2 = 34$). Figure 1 reveals the error bar of the test scores and the PPBPD for the two groups. In comparison with the healthy group, the BPD group had a higher mean score and mean PPBPD on the CDA-BPD. Specifically, the healthy group had a mean CDA-BPD score of 15.76 ($SD = 9.595$) and a mean PPBPD of 0.184 ($SD = 0.355$), while the BPD group had a mean score of 45.12 ($SD = 7.231$) and a mean PPBPD of 0.985 ($SD = 0.084$). In addition, paired-sample t tests were conducted to determine the significance of paired-group differences in mean score and PPBPD of the CDA-BPD. The paired-sample t-test statistic of the mean CDA-BPD score of the two groups was −14.246 ($df = 33$, $p < .001$, and Cohen's $d = 3.330$) and the t-test statistic of the mean PPBPD of the two groups was −12.854 ($df = 33$, $p < .001$ and Cohen's $d = 3.083$). To sum up, there were different CDA-BPD scores and PPBPD between the two groups, and the distributions were also reasonably symmetric within the two groups. These results indicate that the CDA-BPD could effectively separate the healthy group from the BPD group.

Moreover, the content validity was also investigated in this article. Just like the BPQ scale (Poreh et al., 2006), the CDA-BPD also contained two underlying factors: identity/interpersonal and impulsivity. In the CDA-BPD, there were 38 items to measure the identity/interpersonal trait and 17 items to measure impulsivity trait. Furthermore, the identity/interpersonal factor assessed abandonment (S1, included 4 items), relationships (S2, included 6 items), self-image (S3, included 6 items), affective instability (S6, included 8 items), emptiness (S7, included 10 items) and quasi-psychotic states (S9, included 4 items). The impulsivity factor assessed impulsivity (S4, included 4 items), suicide (S5, included 6 items), and intense anger (S8, included 7 items). Beyond that, the CFA results of the CDA-BPD (see section of structural parameter results) also showed that CDA-BPD covered all the domains that were suggested by the BPQ (Poreh et al., 2006).

Sensitivity and specificity

Based on the validation sample, diagnostic accuracy of the CDA-BPD was assessed by the nonparametric measure of area under a ROC curve. Results showed that the area under the ROC curve (AUC) was 0.956 (95% CI [0.901, 0.989]), and the CDA-BPD had high sensitivity of 0.985 and high specificity of 0.853 for identifying the diagnostic accuracy of the CDA-BPD. These results indicate that the CDA-BPD had acceptable power to distinguish healthy individuals and BPD individuals.

Diagnostic score reporting

Table 3 and Figure 2 provide details of the three individuals' nine symptom criteria information based on the *DSM-5* and their PPBPD reports, which is an example to illustrate the unique information provided by the CDM. Although these individuals had the same total score in the MIS-BPD and BPQ, and they were all defined as BPD by the MIS-BPD and the BPQ, their posterior probability of meeting each symptom criterion varied considerably (see Table 3) as a result of individual differences. Based on these probabilities, the PPBPD for each individual could be estimated.

More specifically, individuals A, B and C were all diagnosed as BPD by the CDA-BPD with the PPBPD of 0.987, 0.991, and 1.00 respectively, which was consistent with the diagnosis of the MIS-BPD and BPQ, but with variance in their symptom profiles. As seen in Table 3 and Figure 2, Individual A (female, 20 years old and from the country) had a probability of over .5 on symptoms 1, 2, 3, 6, 7, 9 but not on symptoms 4, 5 and 8; Individual B (male, 21 years old and from the country) had a probability of over .5 on symptoms 1, 2, 4, 5, 6, 8, 9 but not on symptoms 3 and 7; and Individual C (female, 19 years old and from the city) had a probability of over 0.5 for all the symptoms. This detailed information may be useful for individual-specific diagnostic and interventional treatment.

Structural parameter results

The structure of the CDA-BPD was confirmed: χ^2 was 83.01 with $df = 26$, $p < .0001$, $\chi^2/df = 3.19$, NFI = 0.93, NNFI = 0.95, CFI = 0.97, GFI = 0.98, ACFI = 0.97, and RMSEA = 0.045. According to these results, all goodness-of-fit indicators were in the range of acceptability. Figure 3 displays the results of the

https://doi.org/10.1017/prp.2019.14 Published online by Cambridge University Press

Table 2. The 55 selected items of the final CDA-BPD

				Model-fit (item level)			DIF (Female and male)			DIF (Rural and urban)		
Item	g_j	s_j	$Disc_j$	S-X^2	*df*	*p*	Wald statistic	*df*	*p*	Wald statistic	*df*	*p*
1	0.050	0.604	0.346	43.88	42	.392	1.48	2	.476	0.13	2	.936
2	0.124	0.368	0.507	57.68	42	.054	0.17	2	.917	1.27	2	.531
3	0.025	0.589	0.386	58.66	42	.045	3.54	2	.170	6.80	2	.033
6	0.013	0.698	0.289	35.96	42	.732	6.04	2	.049	2.79	2	.248
7	0.042	0.725	0.232	58.30	42	.048	2.50	2	.287	2.20	2	.333
9	0.124	0.396	0.480	43.05	42	.426	5.17	2	.076	1.42	2	.492
11	0.240	0.348	0.412	57.54	42	.056	7.17	2	.028	3.80	2	.150
12	0.198	0.499	0.304	33.80	42	.812	7.96	2	.019	0.48	2	.787
13	0.088	0.610	0.302	35.63	42	.745	0.76	2	.684	3.44	2	.179
14	0.063	0.415	0.522	35.62	42	.746	1.68	2	.431	4.09	2	.130
15	0.094	0.504	0.402	47.82	42	.248	0.20	2	.905	2.05	2	.359
16	0.054	0.506	0.440	42.10	42	.467	5.77	2	.056	0.76	2	.683
18	0.475	0.199	0.327	52.29	42	.133	2.46	2	.292	0.46	2	.793
19	0.060	0.522	0.418	51.16	42	.157	1.12	2	.570	0.64	2	.725
21	0.049	0.467	0.484	35.86	42	.737	1.69	2	.429	0.88	2	.643
22	0.013	0.709	0.279	46.22	42	.302	4.25	2	.119	1.38	2	.501
23	0.185	0.426	0.389	40.00	42	.559	4.47	2	.107	4.22	2	.121
24	0.022	0.522	0.456	39.03	42	.602	0.49	2	.782	7.14	2	.028
27	0.066	0.529	0.405	49.40	42	.202	1.09	2	.580	1.32	2	.516
29	0.010	0.721	0.269	52.22	42	.134	3.42	2	.181	2.43	2	.297
30	0.000	0.711	0.289	29.54	42	.926	0.65	2	.721	4.34	2	.114
31	0.029	0.592	0.379	55.98	42	.073	0.19	2	.911	2.46	2	.292
35	0.032	0.627	0.341	43.73	42	.398	1.59	2	.452	0.26	2	.878
36	0.059	0.504	0.437	58.16	42	.050	1.24	2	.539	3.29	2	.194
37	0.153	0.395	0.452	33.73	42	.815	1.16	2	.560	0.19	2	.910
38	0.000	0.794	0.206	25.07	42	.982	1.02	2	.601	1.59	2	.451
39	0.160	0.476	0.364	41.51	42	.492	0.60	2	.741	0.46	2	.793
40	0.023	0.556	0.421	52.04	42	.138	0.59	2	.743	6.73	2	.035
41	0.073	0.601	0.326	51.63	42	.147	3.79	2	.151	1.04	2	.593
46	0.023	0.771	0.206	53.99	42	.102	7.21	2	.027	1.37	2	.505
47	0.028	0.484	0.488	42.99	42	.428	7.37	2	.025	0.06	2	.969
49	0.048	0.536	0.416	45.21	42	.340	0.84	2	.656	0.25	2	.881
51	0.261	0.412	0.327	41.16	42	.508	3.57	2	.168	0.05	2	.975
54	0.325	0.276	0.400	63.44	42	.018	0.33	2	.847	0.67	2	.716
55	0.134	0.475	0.391	43.84	42	.393	7.49	2	.024	1.11	2	.575
56	0.048	0.601	0.351	45.55	42	.327	3.40	2	.183	1.68	2	.432
60	0.293	0.395	0.312	50.00	42	.185	3.93	2	.140	4.62	2	.099
61	0.033	0.503	0.464	49.84	42	.190	0.27	2	.875	0.85	2	.655
62	0.028	0.514	0.458	53.08	42	.117	2.85	2	.241	3.02	2	.221
63	0.062	0.657	0.280	40.42	42	.540	3.27	2	.195	3.34	2	.188
64	0.008	0.768	0.224	37.50	42	.669	3.46	2	.178	2.56	2	.278
65	0.018	0.633	0.349	58.11	42	.050	1.74	2	.419	0.31	2	.857

(*Continued*)

https://doi.org/10.1017/prp.2019.14 Published online by Cambridge University Press

Table 2. (*Continued*)

				Model-fit (item level)			DIF (Female and male)			DIF (Rural and urban)		
Item	g_j	s_j	$Disc_j$	S-X^2	*df*	*p*	Wald statistic	*df*	*p*	Wald statistic	*df*	*p*
68	0.067	0.400	0.533	31.77	42	.875	2.19	2	.335	3.29	2	.193
69	0.096	0.504	0.400	32.46	42	.855	0.04	2	.979	1.33	2	.514
70	0.399	0.333	0.269	51.55	42	.148	0.10	2	.952	0.48	2	.785
71	0.079	0.539	0.381	49.33	42	.204	7.56	2	.023	2.23	2	.329
72	0.017	0.673	0.310	58.07	42	.051	1.14	2	.566	2.29	2	.318
73	0.010	0.658	0.332	42.36	42	.456	3.19	2	.203	2.21	2	.331
74	0.040	0.622	0.338	57.83	42	.053	0.49	2	.782	2.66	2	.264
75	0.000	0.794	0.206	29.50	42	.927	4.70	2	.096	2.86	2	.239
76	0.046	0.531	0.423	39.28	42	.591	2.06	2	.358	1.17	2	.558
77	0.012	0.630	0.358	40.53	42	.536	1.67	2	.435	1.90	2	.387
78	0.073	0.595	0.332	44.41	42	.370	5.58	2	.061	1.62	2	.444
79	0.283	0.336	0.381	56.29	42	.069	4.82	2	.090	2.89	2	.236
80	0.014	0.632	0.354	43.66	42	.401	3.40	2	.183	4.96	2	.084

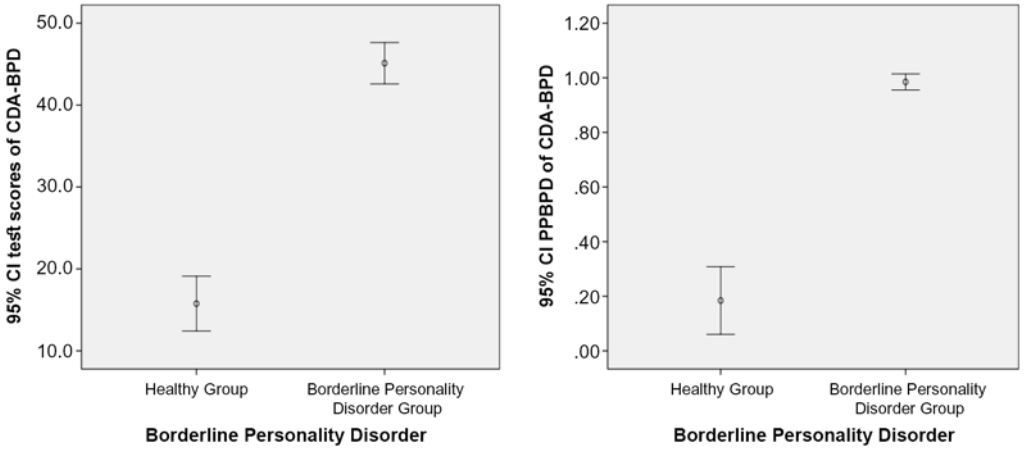

Figure 1. Error bar of the CDA-BPD test scores and PPBPD for validation sample. Note: 95% CI = 95% confidence interval. PPBPD = posterior probability of borderline personality disorder, which was measured based on the CDA-BPD and the diagnostic criteria in the *DSM-5* via CDMs.

Table 3. Individual example estimates

	Individuals		
Symptom criterion	A	B	C
S1	0.911	0.755	0.809
S2	0.944	0.999	0.997
S3	1.000	0.003	0.991
S4	0.200	0.852	0.807
S5	0.085	1.000	0.745
S6	0.914	1.000	1.000
S7	1.000	0.008	0.996
S8	0.043	1.000	1.000
S9	0.995	0.751	0.992
PPBPD	0.987	0.991	1.000

Note: S1–S9 represent nine symptom criteria for BPD defined in the *DSM-5* in Table 1; PPBPD = the posterior probability of borderline personality disorder, which was measured based on the CDA-BPD and the diagnostic criteria in the *DSM-5* via CDMs.

structural model estimation of the CDA-BPD. In Figure 3, the correlation between two latent continuous factors was high at .940, which suggested that the presence or absence of one component would have a significant impact on the existence or absence of the other. The nine symptom criteria had high factor loadings with values typically above 0.60 ($p < .01$). These results indicated that the nine latent criteria of BPD were highly correlated with the two factors, which provided more information for the etiology and was valuable for the diagnosis of BPD.

Additionally, Figure 3 displays the results of the cut-point parameters estimation. These cut-point parameters offered some insights into the percentages of individuals meeting each *DSM-5* criterion in the present sample. To be specific, an estimated 38.9% of respondents satisfied symptom 7 ("Chronic feelings of emptiness"). Furthermore, an estimated 37.6% of respondents met symptom 3 ("Identity disturbance: markedly and persistently unstable self-image or sense of self"). These results suggest that individuals in the current sample who probably had BPD often felt chronic emptiness and had an unstable self-image. The structure results analyzed in the study provided a way to research the

https://doi.org/10.1017/prp.2019.14 Published online by Cambridge University Press

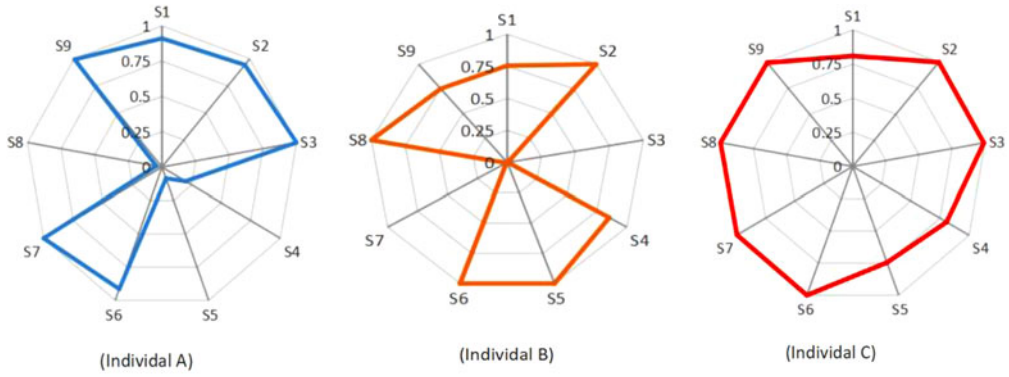

Figure 2. Symptom spectrum of borderline personality disorder for three individuals. Note: S1–S9 represent nine symptom criteria for BPD in the *DSM-5* shown in Table 1.

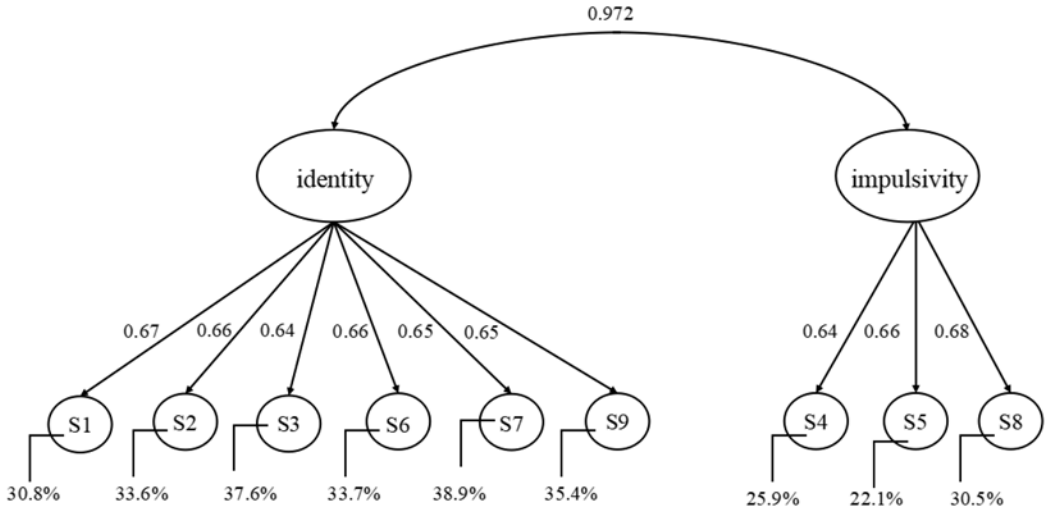

Figure 3. CDA-BPD analysis structural model estimates.

pathological diagnosis and etiological manifestation. The combination of marginal criteria information from the cut-point parameters and pairwise criteria association information from the factor loadings made the two-factor structural model an informative addition to the measurement component of the CDM.

Discussion

The present study reported on the process of retrofitting the traditional borderline personality disorder questionnaire and initial validation of the CDA-BPD. The CDA-BPD was improved under the CDM framework based on the *DSM-5*, and the results of this study suggested that the CDA-BPD was a reliable self-report instrument to evaluate BPD. In particular, the retrofitting process of the CDA-BPD showed an innovative statistical and psychometric approach ready to provide effective formative information at a more fine-grained level than currently offered by the ubiquitous unidimensional instrument.

Compared with the previous researches on BPD assessments, this study had some major implications, which are listed as follows.

Why was the CDA-BPD better than traditional BPD assessments? First, although all the items of the CDA-BPD were directly based on the traditional BPQ questionnaire (Poreh et al., 2006), the 55 items that consisted of the final CDA-BPD were carefully selected according to the statistical indicators of a CDM. In the present study, most items of the CDA-BPD had a strong power to discriminate individuals who present the attributes on an item and are absent the attributes. The $S\text{-}X^2$ statistic of all remaining 55 items were acceptable and indicated a good fit between the estimated G-DINA model and the observed data on item level. The CDA-BPD deleted the items with DIF under the CDM framework so as to remove any bias in the current assessment. Besides, analysis results also showed that the CDA-BPD not only had good reliability, *convergent* validity and content validity, but also had high sensitivity and specificity. These convincing evidences indicate that the CDA-BPD was a high-quality BPD assessment. Second, the retrofitting process of the CDA-BPD was under the CDM framework. The CDM as a combination of psychological assessment and clinical diagnosis can be used not only to evaluate psychometric properties for test and items but also to provide diagnostic information at both the generic level and symptom level of diagnostic information. Based on the above, the CDA-BPD is a high-quality assessment, and compared with the traditional BPD scale, it can also provide detailed cognitive diagnostic information. So, the CDA-BPD can be regarded as an improved assessment tool.

What is the difference between the proposed method of CDM and latent class analysis for the same purposes? Latent class analysis is a statistical method used to identify a set of discrete, mutually exclusive latent classes of individuals based on their responses to a set of observed categorical variables. It can classify individuals from

https://doi.org/10.1017/prp.2019.14 Published online by Cambridge University Press

a heterogeneous population into smaller, relatively homogenous unobserved subgroups. CDMs are special cases of latent class models but with some differences between them. On the one hand, unlike unrestricted latent class models, CDMs apply a hypothesized set of constraints to measure a set of preconceived dichotomous latent variables or dichotomous attributes. These constraints specify the attributes required to positively respond to each item, and in CDMs are reflected in the Q-matrix. Due to such constraints (Q-matrix), the general response characteristics of each individual are known in CDMs. On the other hand, latent class analysis is regarded as the baseline model for a hypothesis model; that is, latent class analysis assumes that there is complete independence between observed variables, then increases the number of potential classes gradually from baseline model and selects the appropriate number of classes (normally, the number of classifications would not be so large) and maximizes the model fit. But in CDMs, the number of classes are fixed and the meaning of class membership for the analysis is also defined when the number of attributes is known (e.g. if there are K attributes, the sample can be divided into 2^K). Future studies may apply unrestricted latent class models to investigate BPD and to compare with the CDM results in the CDA-BPD.

Could the proposed method of retrofitting a traditional BPD scale under the CDM be applied to interval scale variables? As mentioned above, CDMs were latent variable models developed primarily to describe the relationship between observable data (typically in the form of questionnaire responses) and a set of categorical latent variables (typically dichotomous or binary-valued [0/1]). It should be pointed out that the initial observation data and Q matrix of the CDA-BPD were both dichotomous in the present study. But, if researchers want to apply CDMs to investigate interval scale variables, there are still some viable solutions. As the items of the instrument are in the form of a Likert-type scale, researchers could follow Templin and Henson (2006), who converted the partial-credit items to dichotomous items by considering full credit as success (recorded as 1), and the remaining scores as failure (recorded as 0). Another way is that the score of zero does not change but a score of more than zero is converted to 1 (Y.S. Lee, Park, & Taylan, 2011). Note that using the seq-GDINA model (Ma & de la Torre, 2016) directly is an acceptable way to deal with Likert-type scale data. As the attribute of the Q-matrix was polytomous, more detailed information could also be referred to in Chen and de la Torre (2013).

Accordingly, the main contributions of this study were that: (1) For the first time, a traditional BPD questionnaire was retrofitted for diagnostic purposes under the CDM framework. The process of retrofitting in this article could serve as an important step in advancing BPD assessment research when it has been too difficult to develop a cognitive diagnostic BPD instrument due to practical constraints and as researchers want to obtain more accurate and rich information for diagnosis. In terms of the new psychometric method used in this article, the CDM is an effective statistical method for use in psychological research. It may be able to enhance research into other personality disorders, in which the dichotomous attribute nature of the CDM can help researchers understand the detailed components of personality traits. (2) As an improved BPD assessment in this study, the CDA-BPD could provide both general diagnostic information (e.g. BPD vs. non-BPD) and detailed information about how each individual meets the definition for diagnosis. The criterion-level information gives insight into tailoring individual-specific treatments for BPD, potentially increasing these treatments' effectiveness. (3) In the study, we investigated the structure of underlying personality factors in BPD to provide diagnostic information for each criterion. Using the factor structure estimation from a CDM, we were able to provide avenues to research the evaluation and etiological manifestations of BPD.

Although the results were promising, there are also several limitations to acknowledge. First, the participants in this study were college students who were from a narrow population distribution. The etiology of BPD may be different for young adults. Second, in the validation sample, the BPD group was recruited from students who were defined as BPD by the MIS-BPD and PDQ-4-BPD; however, it is more reasonable that the BPD group should be strictly regarded clinically as having BPD. Third, the CDM used in this study was more intricate than other methods, and it needed to estimate many parameters. Any interested reader could refer to Ma, Iaconangelo, and de la Torre (2016) or Sorrel, Abad, Olea, de la Torre, and Barrada, (2017) to learn how to overcome the limitations. Finally, while retrofitting under the CDM framework showed a feasible method to gain more effective information from existing BPD assessments, researchers should be careful about using this method if the instruments do not specify a correct Q-matrix.

Conclusion

This article introduced the process of retrofitting the BPD assessment under the CDM framework, and the results provided sufficient evidence for the CDA-BPD as an improved BPD assessment. The CDA-BPD may also be a valuable tool to provide rich diagnosis information that the traditional BPD assessment could not. With this assessment, the researchers could estimate each individual's symptom profile and the PPBPD according to the *DSM-5*. This rich diagnostic information could potentially increase clinical treatments' effectiveness.

Financial Support. This study was funded by National Natural Science Foundation of China (31660278, 31760288).

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Aggen S.H., Neale M.C., Røysamb E., Reichborn-Kjennerud T. and Kendler K.S. (2009). A psychometric evaluation of the DSM-IV borderline personality disorder criteria: Age and sex moderation of criterion functioning. *Psychological Medicine*, **39**, 1967–1978.

Bach B., Sellbom M., Bo S. and Simonsen E. (2016). Utility of DSM-5 section III personality traits in differentiating borderline personality disorder from comparison groups. *European Psychiatry*, **37**, 22–27.

Browne M.W. and Cudeck R. (1993). Alternative ways of assessing model fit. *Sage Focus Editions*, **154**, 136–136.

Byrne B.M. (2012). A primer of LISREL: Basic applications and programming for confirmatory factor analytic models. New York: Springer Verlag.

Chen H. and Chen J. (2016). Retrofitting non-cognitive-diagnostic reading assessment under the generalized DINA model framework. *Language Assessment Quarterly*, **13**, 218–230.

Chen H., Zhong J. and Liu Y.X. (2011). Application of Mclean screening instrument for borderline personality disorder in Chinese psychiatric samples. *Chinese Journal of Clinical Psychology*, **19**, 594–595.

https://doi.org/10.1017/prp.2019.14 Published online by Cambridge University Press

Chen J. and de la Torre J. (2013). A general cognitive diagnosis model for expert-defined polytomous attributes. *Applied Psychological Measurement*, **37**, 419–437.

Conway C., Hammen C. and Brennan P. (2012). A comparison of latent class, latent trait, and factor mixture models of DSM-IV borderline personality disorder criteria in a community setting: Implications for DSM-5. *Journal of Personality Disorders*, **26**, 793–803.

Cui Y., Gierl M.J. and Chang H.H. (2012). Estimating classification consistency and accuracy for cognitive diagnostic assessment. *Journal of Educational Measurement*, **49**, 19–38.

de la Torre J. (2008). An empirically based method of Q-matrix validation for the DINA model: Development and applications. *Journal of Educational Measurement*, **45**, 343–362.

de la Torre J. (2009). A cognitive diagnosis model for cognitively based multiple-choice options. *Applied Psychological Measurement*, **33**, 163–183.

de la Torre J. (2011). The generalized DINA model framework. *Psychometrika*, **76**, 179–199.

de la Torre J. and Chen J. (2011, April). Estimating different reduced cognitive diagnosis models using a general framework. *Paper presented at the annual meeting of the National Council on Measurement in Education*, New Orleans, LA.

de la Torre J. and Chiu C.Y. (2016). General method of empirical Q-matrix validation. *Psychometrika*, **81**, 253–273.

de la Torre J. and Minchen N. (2014). Cognitively diagnostic assessments and the cognitive diagnosis model framework. *Psicología Educativa*, **20**, 89–97.

de la Torre J., van der Ark L.A. and Rossi G. (2018). Analysis of clinical data from a cognitive diagnosis modeling framework. *Measurement and Evaluation in Counseling and Development*, **51**, 281–296.

Ferrer M., Andión Ó., Calvo N., Hörz S., Fischer-Kern M. and Kapusta N.D. (2018). Clinical components of borderline personality disorder and personality functioning. *Psychopathology*, **51**, 35–46.

Fonseca-Pedrero E., Paino M., Lemos-Giráldez S., Sierra-Baigrie S. and Muñiz J. (2011). Measurement invariance of the Schizotypal Personality Questionnaire-Brief across gender and age. *Psychiatry Research*, **190**, 309–315.

Hartz S.M. (2002). A Bayesian framework for the unified model for assessing cognitive abilities: Blending theory with practicality. *American Journal of Gastroenterology*, **95**, 906–909.

Hou L., de la Torre J. and Nandakumar R. (2014). Differential item functioning assessment in cognitive diagnostic modeling: Application of the Wald test to investigate DIF in the DINA model. *Journal of Educational Measurement*, **51**, 98–125.

Huang J., Napolitano Li A., Wu J., Yang Y., Xi Y., Li Y. and Li K. (2014). Childhood experiences of parental rearing patterns reported by Chinese patients with borderline personality disorder. *International Journal of Psychology*, **49**, 38–45.

Hu L.T. and Bentler P.M. (1999). Cut-off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, **6**, 1–55.

Hyler S.E. (1994). *PDQ-4 and PDQ-4+ instructions for use*. New York: New York State Psychiatric Institute.

Iacono W.G. (2013). Looking forward to the new personality disorder classification for DSM-5. *Personality Disorders: Theory, Research, and Treatment*, **4**, 284–285.

Jaeger J., Tatsuoka C., Berns S.M. and Varadi F. (2006). Distinguishing neurocognitive functions in schizophrenia using partially ordered classification models. *Schizophrenia Bulletin*, **32**, 679–691.

Jang E.E. (2009). Cognitive diagnostic assessment of L2 reading comprehension ability: Validity arguments for fusion model application to LanguEdge assessment. *Language Testing*, **26**, 31–73.

Junker B.W. and Sijtsma K. (2001). Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Applied Psychological Measurement*, **25**, 258–272.

Joreskog K. and Sorbom D. (1996). *LISREL 8: User's reference guide*. Chicago, IL: Scientific Software International.

Kröger C., Huget F. and Röpke S. (2011). Diagnostic accuracy of the McLean Screening Instrument for borderline personality disorder in an inpatient sample who seek a disorder-specific treatment. *PPmP-Psychotherapie Psychosomatik Medizinische Psychologie*, **61**, 481.

Lee Y.S., Park Y.S. and Taylan D. (2011). A cognitive diagnostic modeling of attribute mastery in Massachusetts, Minnesota, and the US national sample using the TIMSS 2007. *International Journal of Testing*, **11**, 144–177.

Lee Y.W. and Sawaki Y. (2009). Application of three cognitive diagnosis models to ESL reading and listening assessments. *Language Assessment Quarterly*, **6**, 239–263.

Liu R., Huggins-Manley A.C. and Bulut O. (2018). Retrofitting diagnostic classification models to responses from IRT-based assessment forms. *Educational and Psychological Measurement*, **78**, 357–383.

Ma W. and de la Torre J. (2016). GDINA: The Generalized DINA Model Framework. R package version 0.13.0. Retrieved from http://CRAN.Rproject.org/package=GDINA

Ma W., Iaconangelo C. and de la Torre J. (2016). Model similarity, model selection, and attribute classification. *Applied Psychological Measurement*, **40**, 200–217.

Miller J.D., Morse J.Q., Nolf K., Stepp S.D. and Pilkonis P.A. (2012). Can DSM-IV borderline personality disorder be diagnosed via dimensional personality traits? Implications for the DSM-5 personality disorder proposal. *Journal of Abnormal Psychology*, **121**, 944–950.

Mullins-Sweatt S.N., Edmundson M., Sauer-Zavala S., Lynam D.R., Miller J.D. and Widiger T.A. (2012). Five-factor measure of borderline personality traits. *Journal of Personality Assessment*, **94**, 475–487.

Orlando M. and Thissen D. (2003). Further investigation of the performance of S-X2: An item fit index for use with dichotomous item response theory models. *Applied Psychological Measurement*, **27**, 289–298.

Poreh A.M., Rawlings D., Claridge G., Freeman J.L., Faulkner C. and Shelton C. (2006). The BPQ: A scale for the assessment of borderline personality based on DSM-IV criteria. *Journal of Personality Disorders*, **20**, 247–260.

Roberts M.R. and Gierl M.J. (2010). Developing score reports for cognitive diagnostic assessments. *Educational Measurement: Issues and Practice*, **29**, 25–38.

Sorrel M.A., Abad F.J., Olea J., de la Torre J. and Barrada J.R. (2017). Inferential item-fit evaluation in cognitive diagnosis modeling. *Applied psychological measurement*, **41**, 614–631.

Swets J.A. (1988). Measuring the accuracy of diagnostic systems. *Science*, **240**, 1285–1293.

Tatsuoka K.K. (1990). Toward an integration of item-response theory and cognitive error diagnosis. In N. Frederiksen, R. Glaser, A. Lesgold, and M.G. Shafto (Eds.), *Diagnostic monitoring of skill and knowledge acquisition* (pp. 453–488). Hillsdale, NJ: Lawrence Erlbaum Associates.

Templin J. and Henson R.A. (2006). Measurement of psychological disorders using cognitive diagnosis models. *Psychological Methods*, **11**, 287–305.

Templin J. and Henson R.A. (2010). *Diagnostic measurement: Theory, methods, and applications*. New York, NY: Guilford Press.

Templin J. and Bradshaw L. (2013). Measuring the reliability of diagnostic classification model examinee estimates. *Journal of Classification*, **30**, 251–275.

Tu D., Gao X., Wang D. and Cai Y. (2017). A new measurement of internet addiction using diagnostic classification models. *Frontiers in Psychology*, **8**, 1768–1776.

Wang Y.Y., Liang Y.J. and Zhong J. (2008). The adaptation of McLean screening instrument of for borderline personality disorder among Chinese college students. *Chinese Journal of Clinical Psychology*, **16**, 258–260.

Yang J., McCrae R.R., Costa P.T., Yao S., Dai X., Cai T. and Gao B. (2000). The cross-cultural generalizability of Axis-II constructs: An evaluation of two personality disorder assessment instruments in the People's Republic of China. *Journal of Personality Disorders*, **14**, 249–263.

Yi Y.S. (2017). In search of optimal cognitive diagnostic model(s) for ESL Grammar test data. *Applied Measurement in Education*, **30**, 123–144.

Zanarini M.C., Vujanovic A.A., Parachini E.A., Boulanger J.L., Frankenburg F.R. and Hennen J. (2003). A screening measure for BPD: The McLean screening instrument for borderline personality disorder (MSI-BPD). *Journal of Personality Disorders*, **17**, 568–573.

Zanarini M.C. (2003). Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD): A continuous measure of DSM-IV borderline psychopathology. *Journal of Personality Disorders*, **17**, 233–242.

https://doi.org/10.1017/prp.2019.14 Published online by Cambridge University Press

Appendix A

The initial Q-matrix of the BPQ scale

	Q-matrix								
Item	S1	S2	S3	S4	S5	S6	S7	S8	S9
1#	0	0	0	1	0	0	0	0	0
2#	0	0	0	0	0	1	0	0	0
3#	1	0	0	0	0	0	0	0	0
4	0	1	0	0	0	0	0	0	0
5	0	0	1	0	0	0	0	0	0
6#	0	0	0	0	1	0	0	0	0
7#	0	0	0	0	0	0	1	0	0
8	0	0	0	0	0	0	0	1	0
9#	0	0	0	0	0	0	0	0	1
10	0	0	0	1	0	0	0	0	0
11#	0	0	0	0	0	1	0	0	0
12#	1	0	0	0	0	0	0	0	0
13#	0	1	0	0	0	0	0	0	0
14#	0	0	1	0	0	0	0	0	0
15#	0	0	0	0	1	0	0	0	0
16#	0	0	0	0	0	0	1	0	0
17	0	0	0	0	0	0	0	1	0
18#	0	0	0	0	0	0	0	0	1
19#	0	0	0	0	0	1	0	0	0
20	1	0	0	0	0	0	0	0	0
21#	0	1	0	0	0	0	0	0	0
22#	0	0	0	0	1	0	0	0	0
23#	0	0	0	0	0	0	1	0	0
24#	0	0	0	0	0	0	0	1	0
25	0	0	0	0	0	0	0	0	1
26	0	0	0	1	0	0	0	0	0
27#	0	0	0	0	0	1	0	0	0
28	1	0	0	0	0	0	0	0	0
29#	0	1	0	0	0	0	0	0	0
30#	0	0	0	0	1	0	0	0	0
31#	0	0	0	0	0	0	1	0	0
32	0	0	0	0	0	0	0	1	0
33	0	0	0	0	0	0	0	0	1
34	0	0	0	1	0	0	0	0	0
35#	0	0	0	0	0	1	0	0	0
36#	0	1	0	0	0	0	0	0	0
37#	0	0	1	0	0	0	0	0	0
38#	0	0	0	0	1	0	0	0	0
39#	0	0	0	0	0	0	1	0	0
40#	0	0	0	0	0	0	0	1	0
41#	0	0	0	0	0	0	0	0	1
42	0	0	0	1	0	0	0	0	0
43	0	0	0	0	0	1	0	0	0
44	1	0	0	0	0	0	0	0	0
45	0	1	0	0	0	0	0	0	0
46#	0	0	1	0	0	0	0	0	0
47#	0	0	0	0	0	0	0	1	0
48	0	0	0	0	0	0	0	0	1
49#	0	0	0	0	0	1	0	0	0
50	1	0	0	0	0	0	0	0	0
51#	0	1	0	0	0	0	0	0	0
52	0	0	1	0	0	0	0	0	0
53	0	0	0	0	1	0	0	0	0
54#	0	0	0	0	0	0	1	0	0
55#	0	0	0	0	0	0	0	1	0
56#	0	0	0	0	0	0	0	0	1
57	0	0	0	1	0	0	0	0	0
58	0	0	0	0	0	1	0	0	0
59	1	0	0	0	0	0	0	0	0
60#	0	1	0	0	0	0	0	0	0
61#	0	0	1	0	0	0	0	0	0
62#	0	0	0	0	0	0	1	0	0
63#	0	0	0	0	0	0	0	1	0
64#	0	0	0	1	0	0	0	0	0
65#	0	0	0	0	0	1	0	0	0
66	1	0	0	0	0	0	0	0	0
67	0	0	1	0	0	0	0	0	0
68#	0	0	0	1	0	0	0	0	0
69#	0	0	0	0	0	0	1	0	0
70#	0	0	1	0	0	0	0	0	0
71#	0	0	0	1	0	0	0	0	0
72#	0	0	0	0	0	1	0	0	0
73#	1	0	0	0	0	0	0	0	0
74#	0	0	1	0	0	0	0	0	0
75#	0	0	0	0	1	0	0	0	0
76#	0	0	0	0	0	0	1	0	0
77#	0	0	0	0	0	0	0	1	0
78#	1	0	0	0	0	0	0	0	0
79#	0	0	0	0	0	0	1	0	0
80#	0	0	0	0	0	0	0	1	0

Note: S1–S9 represented nine symptom criteria for borderline personality disorder defined in DSM-5. Items marked with an asterisk (#) were retained in the final 55-items of the CDA-BPD.

https://doi.org/10.1017/prp.2019.14 Published online by Cambridge University Press

Appendix B

Note that items marked with an asterisk (*) are reverse scored.

The English version of the final CDA-BPD

1. I often do things without thinking them through.
2. I often become depressed or anxious "out of the blue".
3. People often leave me.
6. I have threatened to hurt myself in the past.
7. I do not believe that I have the skills to do anything with my life.
9. Sometimes I feel like I am not real.
11. I sometimes feel anxious or irritable and become sad a few hours later.
12. When people close to me die or leave me, I feel abandoned.
13. I often exaggerate the potential of friendships only to find out later that they will not work out.
14. If I were more like other people, I would feel better about myself.
15. I have deliberately tried to hurt myself without trying to kill myself.
16. In general, my life is pretty boring.
18. People are sometimes out to get me.
19. My friends have told me that my mood changes very quickly.
21. People who seem trustworthy often disappoint me.
22. I have made a suicide attempt in the past.
23. I often feel like I have nothing to offer others.
24. I have trouble controlling my temper.
27. My mood frequently alternates throughout the day between happiness, anger, anxiety, and depression.
29. My friends often disappoint me.
30. I have cut myself on purpose.
31. I often feel lonely and deserted.
35. I sometimes feel very sad, but this feeling can change quickly.
36. People often let me down.
37. I wish I could be more like some of my friends.
38. I used to try to hurt myself to get attention.
39. I am often different with different people in different situations so that sometimes I am not sure who I am.
40. I easily become irritated by others.
41. Sometimes I can actually hear what other people are thinking.
46. I feel that people would not like me if they really knew me well.
47. I get angry easily.
49. I sometimes feel very happy but this feeling can change quickly.
51. The relationships with people I care about have lots of ups and downs.
54. I rarely feel lonely.*
55. I often find that the littlest things make me angry.
56. Sometimes I can't tell between what is real and what I have imagined.
60. My friends are always there when I need them.*
61. I wish I were someone else.
62. I feel like my life is not interesting.
63. When I am angry, I often hit objects and break things.
64. I often receive speeding tickets.
65. I often feel like I am on an emotional 'roller-coaster.'
68. I often do things impulsively.
69. My life is without purpose.
70. I am not sure what I want to do in the future.
71. At times I eat so much food that I am in pain or have to force myself to throw up.
72. People tell me that I am a moody person.
73. The people I love often leave me.
74. In social situations, I often feel that others will see through me and realize that I don't have much to offer.
75. I have been in the hospital for trying to harm myself.
76. I often feel empty inside.
77. Others often make me angry.
78. I often become frantic when I think that someone I care about will leave me.
79. I am often confused about my long term goals.
80. Others say I'm quick tempered.

The Chinese version of the final CDA-BPD

1. 我做事常常不考虑后果。
2. 我经常感到情绪低落或焦虑不安。
3. 人们经常离开我。
6. 在过去我曾扬言要伤害自己。
7. 我不相信我有能力做好任何与我生活有关的事情。
9. 有时我觉得自己不是真实的。
11. 我有时会感到焦虑或烦躁并且经过几个小时后便感到悲伤
12. 当身边的人死去或离开时我感觉自己被遗弃了。
13. 我常常会夸大友谊的作用，后来发现他们其实并不重要。
14. 如果我更像其他人，我的自我感觉会更好些。
15. 我曾故意伤害自己但没有想过要自杀。
16. 总的来说我的生活十分无聊。
18. 有时别人无法理解我。
19. 我的朋友曾说我的情绪变化无常。
21. 那些看起来值得信任的人却常常使我感到失望。
22. 我曾试图自杀。
23. 我常常觉得自己没有什么东西可以给别人。
24. 我很难控制自己的脾气。
27. 我的情绪在一天之内会在幸福、愤怒、焦虑和抑郁之间频繁交替着。
29. 我的朋友经常使我失望。
30. 我曾故意割伤自己。
31. 我经常感到孤独和被遗弃。
35. 我有时候会感到非常悲伤但这种变化很快。
36. 别人经常让我感到情绪低落。
37. 我希望我能更像我的一些朋友。
38. 我过去经常试图伤害自己以引起别人的注意。
39. 在不同的场合与不同的人相处时我经常表现的不同，所以有时我不确定自己到底是谁。
40. 我很容易被别人激怒。
41. 有时我能够听到别人正在想什么。
46. 当人们了解我之后他们就不会再喜欢我。
47. 我很容易生气。
49. 我有时候会感到非常快乐但这种感觉变化很快。
51. 我和我在乎的人之间的关系经常会有许多起伏。
54. 我很少会感到孤独。(*)
55. 我经常发现一些鸡毛蒜皮的事都会使我很生气。
56. 有时我无法分清什么是事实什么是我想象的。
60. 当我需要帮助时，我的朋友一定会挺身而出。(*)
61. 我希望我是其他人。
62. 我觉得我的生活很没趣。
63. 当我生气时我会摔打东西。
64. 我经常收到超速罚单。

https://doi.org/10.1017/prp.2019.14 Published online by Cambridge University Press

65. 我经常感到自己的情绪变化很大就如同坐过山车一样。
68. 我经常会很冲动地做事情。
69. 我的生活没有目标。
70. 我还不确定未来想做什么。
71. 我有时候会吃很多东西以至于让自己感到很痛苦或者要强迫自己把它丢掉。
72. 别人说我是个喜怒无常的人。
73. 我爱的人总是会离开我。
74. 在社交场合中我常常觉得别人会看穿我并且会觉得我一无所有。
75. 我曾因自残而住过医院。
76. 我经常感到内心很空虚。
77. 别人经常惹我生气。
78. 每当我想到在乎的人会离开时我就会变得很疯狂。
79. 我经常对自己未来的目标感到迷茫。
80. 别人说我很容易动怒。

https://doi.org/10.1017/prp.2019.14 Published online by Cambridge University Press