Application of the Rasch model in measuring personality disorders

Aplicação do modelo de Rasch na medida de transtornos da personalidade

Lucas de Francisco Carvalho,[1] Ricardo Primi,[2] Gregory J. Meyer[3]

Abstract

Objectives: To describe item and person parameters obtained with the Rasch model, one of the item response theory models, in the assessment of personality disorders based on Millon's theory. **Method:** A total of 350 people participated in the study. Age ranged from 18 to 67 years (mean ± standard deviation = 27.02±10.13), and 71.7% of the participants (n = 251) were female. Of the 350 individuals, 21.1% (n = 74) answered affirmatively about being under psychiatric treatment and taking psychiatric medications. The Personality Disorders Dimensional Inventory (PDDI), an instrument designed to assess personality disorders according to Millon's theory, was applied to all participants. Data were analyzed using the Rasch model. **Results:** Overall, analysis with the Rasch model revealed that the PDDI has adequate psychometric properties for the assessment of personality disorders. **Conclusion:** Among the contributions of item response theory models for clinical instruments, the Rasch person-item map deserves to be highlighted as a successful attempt to improve the understanding of clinical scores obtained in response to particular test items.

Keywords: Health evaluation, diagnosis, validation studies, psychiatric status rating scales.

Resumo

Objetivos: Descrever os parâmetros de itens e pessoas obtidos com o modelo de Rasch, um dos modelos da teoria de resposta ao item, na avaliação de transtornos da personalidade de acordo com a teoria de Millon. **Método:** Participaram do estudo 350 pessoas. A idade variou de 18 a 67 anos (média ± desvio padrão = 27.02±10.13), e 71.7% (n = 251) eram do sexo feminino. Dos 350 participantes, 21.1% (n = 74) responderam afirmativamente sobre estarem em tratamento psiquiátrico e utilizarem medicamento psiquiátrico. O Inventário Dimensional de Transtornos da Personalidade (IDTP), destinado à avaliação dos transtornos da personalidade com base na teoria de Millon, foi aplicado a todos os participantes. Os dados foram analisados por meio do modelo de Rasch. **Resultados:** De modo geral, os resultados da análise por meio do modelo de Rasch revelaram que o IDTP tem parâmetros psicométricos adequados para a avaliação de transtornos da personalidade. **Conclusão:** Entre as contribuições da teoria de resposta ao item para instrumentos clínicos, o mapa de pessoas-itens do modelo Rasch merece destaque enquanto tentativa bem-sucedida de aumentar a compreensão das pontuações clínicas obtidas em resposta a determinadas questões de um teste.

Descritores: Avaliação em saúde, diagnóstico, estudos de validação, escalas de graduação psiquiátrica.

Introduction

Personality is understood as a continuum that can be characterized, on the one hand, by modes of functioning that are successful in handling obstacles of everyday life or, on the other hand, by maladaptive functioning.[1] When the psychological functioning of an individual is maladaptive, characterized by shortcomings, dissonance, and conflicts in their ability to deal with the environment, and longstanding, this pattern is regarded as a personality disorder.[2,3]

In Millon's theory, personality disorders are treated as theoretical constructs used to represent different styles or patterns of maladaptive personality functions in relation to one's environment.[4] Millon proposes three spheres that are based on evolutionary principles, named evolutionary stages: aims of existence, modes of adaptation, and strategies of replication.[1,5-8]

The first stage, aims of existence, refers to the tendency of the adapted organism to express mechanisms for increasing and preserving quality of life. Accordingly, the bipolarity present at this stage (pleasure-pain) is characterized, on one side, by pleasure seeking (increasing quality of life) and, on the other side, by pain avoidance (preserving quality of life). The next stage, modes of adaptation, is related to the typical ways adopted to render possible any necessary exchanges between the organism and the environment. The bipolar representative of this stage (active-passive) is the tendency to modify one's environmental niche (active), in contrast with the tendency to adapt to the environment (passive). Finally, the evolutionary stage of strategies of replication refers to the strategies developed by individuals to overcome the limitations of their own existence. As a result, the bipolarity (self-other) here is related to the perpetuation of self vs. the focus on the care and protection of the offspring (other). As a result of the patterns observed for each of these polarities, 15 personality styles are derived.[1,3,4,6] Figure 1 shows these styles according to the polarities of each stage.

Figure 1 summarizes the polarities and personality disorder constructs of Millon's theory based on current literature.[2-4] As can be observed, different personality styles show imbalances in different bipolarities, with a clear trend towards one of the poles. Furthermore, they may be deficits in some bipolarities, signaling difficulty in experiencing or dealing with the poles of some evolutionary stage. In addition, some bipolarities may show reversal or conflicts. Reversal refers to a change in the polarity of a particular evolutionary stage, whereas conflict relates to an ambivalence between the two poles of the stage.

Based on Millon's theory, the Millon Clinical Multiaxial Inventory was developed (MCMI[9]) to assesses personality disorders and other mental illnesses. The latest, third version of the instrument (MCMI-III) contains 14 scales for the assessment of personality disorders. Several studies in the literature have described the psychometric properties of the MCMI-III.[5,9-13]

Despite the incontrovertible importance of the assessment of personality disorders in clinical practice,[8,14,15] few studies in this area have been conducted in Brazil.[16-18] Both nationally and internationally, most of the instruments available for assessing personality disorders have been developed based on classical test theory (CTT). However, over the past few years, it has become increasingly common to find studies, especially in the international literature,[19,20] using an alternative mathematical model to evaluate and guide scale development, namely, item response theory (IRT).[21,22]

The IRT has emerged from criticisms to the classical model. It proposes a mathematical model to represent the testing situation, in which one person answers a set of items. The more intense a given characteristic in the person, the greater the likelihood of agreement with a statement that measures this characteristic. Conversely, the less intense the feature, the smaller the probability that the person will agree.[23] So, the likelihood of choosing a particular answer varies with the degree to which a given characteristic (θ, called theta) is present or not
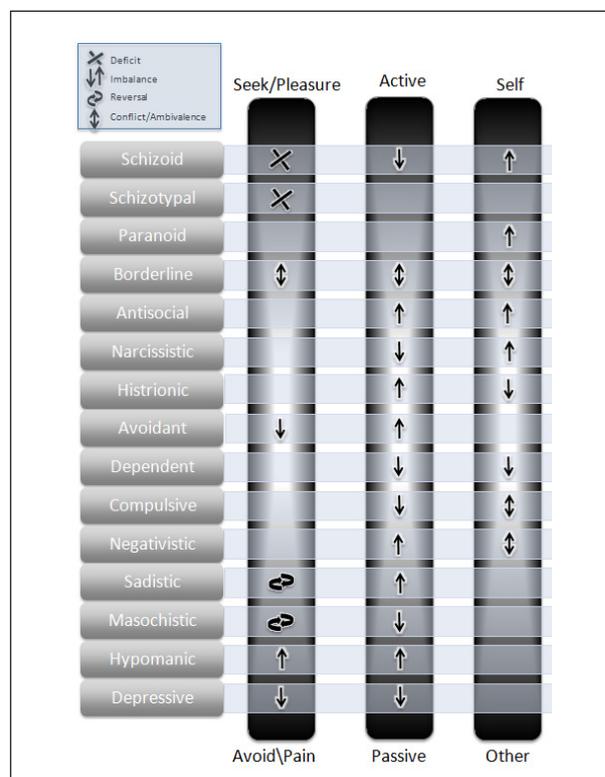
Figure 1 – Styles derived from Millon's theory

in the respondent. There are several models based on IRT, but the Rasch model stands out because of its simplicity and measurement properties.[24] This model parameterizes items according to their intensity while measuring a latent trait; therefore, it has been named one-parameter Rasch model.

For tests using Likert scales, two alternative models are available, derived as extensions and further developments of the Rasch model: the rating scale and the partial credit model,[25,26] also referred to as Andrich's rating scale model and Masters' partial credit model, respectively.[27] Both models conceive the relationship between responses to an item and trait levels (i.e., theta), assuming that each increase in the response scale (e.g., 1 to 5 on a Likert scale) will result in a cumulative step toward higher levels of a latent trait. The basic difference between the two models is that rating scales assume that changes in Likert scores are constant and equal for all items; in the partial credit model, in turn, this condition is relaxed, and different distances between the Likert scores of different items may be considered. In the present study, the rating scale model was used because it is considered to be more appropriate for Likert scales.[28]

The item characteristic curve (ICC) is a trace line generated by the rating scale that expresses the relationship between theta and the probability of a given specific response on a Likert scale. The ICC for the rating scale model is given by the following equation[24]:

$$P_{ijx}\left(\theta_j\right)=\frac{\exp\left(\sum_{j=0}^{x}\left[\theta_j-\left(\lambda_k+b_i\right)\right]\right)}{\sum_{x=0}^{m}\exp\left(\sum_{j=0}^{x}\left[\theta_j-\left(\lambda_k+b_i\right)\right]\right)}$$

where $P_{ijx}\left(\theta_j\right)$ indicates the probability of subject j scoring x on item i.

Item scores have the notation $x = 0, \dots m$ (for $m + 1$ response categories), and also:

$$\sum_{j=0}^{0}\left[\theta_j-\left(\lambda_k+b_i\right)\right]=0$$

The parameter is the measure of person j in the latent dimension theta measured by the item i. Exp(x) signifies the exponentiation of the natural base, @ 2.72, to the x power. The parameter is the threshold of transition (or intersection) between the different categories of the scale. A 4-point Likert scale (e.g., 1, 2, 3, and 4) will have three thresholds representing transition, i.e., between categories 1 and 2 (λ_1), between 2 and 3 (λ_2), and between 3 and 4 (λ_3). These points represent dimensional levels in the transition from one category (e.g., 2) to the next highest category (3) in terms of the probability of selecting these categories. At

the exact point of the threshold (midline), the chances of choosing one or the other category are equal. Above the threshold, the top category becomes more likely to be chosen; below the threshold, the lower category becomes more likely. Therefore, the distance between the thresholds determines the intervals in the latent dimension associated with the probability of a particular response category to occur. A distinctive feature of the rating scale model is that these scalar intervals between categories are relatively similar for all items. The parameter represents the location of item i, or the average intensity of the thresholds of an item. Items that represent extremes in the latent dimension are associated with high average thresholds because their thresholds are all located on the most intense theta levels.

In the recent literature, it is possible to find research using IRT models in the development of personality assessment tests for use in psychological and psychiatric clinical settings.[19,20,28-33] Less frequent, though, are studies that capitalize on the advantage of IRT in its ability to overcome a problem that has become known in the social sciences as arbitrary metrics.[19] Usually, psychological scales are interpreted based on normative references, which attribute meaning to a test score by comparing a person's standing relative to a normative group. Although this information is important, it does not directly inform users about the meaning of a measure per se, or, for instance, what it means to move up or down 2 points in the scale (the arbitrary metrics problem).[34]

In this sense, one advantage of Rasch type models resides in the use of a procedure called item-referenced meaning.[35] This procedure is an alternative way to attribute meaning by linking test scores with expected response patterns, allowing a more qualitatively derived meaning of the scale.[21,36-38] The main objective of the present study was to apply the Rasch model (describing item and person parameters) to a personality disorder test based on Millon's theory, in order to illustrate the advantages of this procedure in scale interpretations.

Method

Participants

A total of 350 people participated in the study. Age ranged between 18 and 67 years (mean ± standard deviation = 27.02±10.13); 71.7% (n = 251) were female. Of the 350 participants, 290 were undergraduate students of psychology (n = 158), architecture (n = 68), and administration (n = 64) at a town in the Brazilian state of São Paulo. The remaining 60 participants were

patients of a psychiatric clinic located in a rural area of São Paulo state. Of the total sample, 21.1% (n = 74) answered affirmatively about being under psychiatric treatment and taking psychiatric medications.

Materials

In accordance with the objectives of this study, the Personality Disorders Dimensional Inventory (PDDI[17]) was administered to all participants. The PDDI was designed to assess personality disorders based on Millon's theory. It comprises 100 items divided into 15 distinct scales (one for each personality style): Depressive (9 items), Schizoid (8 items), Borderline (7 items), Paranoid (6 items), Sadistic (6 items), Compulsive (5 items), Masochistic (7 items), Antisocial (7 items), Dependent (6 items), Histrionic (5 items), Negativistic (6 items), Schizotypal (5 items), Avoidant (7 items), Narcissistic (8 items), and Validity and Social Desirability/Scale XV (8 items). The first 14 scales are sets of items devoted to the assessment of personality disorders, whereas the items comprising Scale XV aim to determine the respondents' motivation to answer the test honestly, as well as their level of social desirability. Scale XV was not analyzed in the present study, once our focus was specifically on the assessment personality disorders. Moreover, it is important to note that one of the styles/scales proposed by Millon, the Hypomanic style, is not represented in the PDDI as it is not present in Millon's English-language scale.[8,9] Each item is answered according to a 4-point Likert scale ranging from 1, "does not describe me," to 4, "describes me extremely or completely." The time necessary to complete the instrument is approximately 20 minutes.

The psychometric properties of the PDDI, such as validity and reliability, have been reported in several studies.[17,39,40] A factor analytic study led by Carvalho[17,39] found empirical support for 12 of the 14 scales theoretically expected to be present, indicating at least a partially adequate factor structure of the instrument. The reliability of scales ranged between 0.65 and 0.89, and only the Narcissistic scale (alpha = 0.65) showed a coefficient of less than 0.71. In another study whose purpose was to assess personality characteristics in patients with chronic pain,[40] two PDDI scales proved to be good predictors of chronic pain: the Histrionic scale (positive correlation) and the Sadistic scale (negative correlation). Moreover, a receiver operating characteristic (ROC) curve revealed that both scales had a sensitivity of 83.3% and a specificity of 31% for the assessment of chronic pain, suggesting high sensitivity (few false negative results) and poor specificity (many false positives).

Procedures and data analysis

The study was revised and approved by the Ethics Committee of Universidade São Francisco, São Paulo, Brazil (protocol no. CAAE 0144.0.142.000.07). The PDDI was administered to all participants after they signed an informed consent form.

One of the authors was always present during instrument administration, available to clarify any possible doubts. Participants completed the PDDI in a classroom of the university or at the waiting room of the psychiatric clinic. Administration was always in groups of patients, and the time required was approximately 25 minutes.

The rating scale model was used, combined with the Winsteps software.[41] Item and subject model parameters were calibrated using the joint maximum likelihood estimation method available on Winsteps. Item difficulties (average thresholds) were set at zero in order to identify scale metrics. Several statistical analyses were used to analyze model fit, including descriptive statistics for subject parameters (theta) and item parameters; model fit indices (infit and outfit); reliability indices; category probability curves; item threshold values; and, most importantly, item map qualitative analysis. Due to scape constraints, detailed analyses will be described for one scale only (Paranoid). All other scales should follow the same principles of analysis.

Results and discussion

Table 1 presents summarized descriptive statistics about the number of items answered in each of the PDDI scales, the latent trait level (theta) of respondents, standard error, and fit indices (infit and outfit). The average level of latent traits found in our sample suggests that, overall, items tended to be difficult for our participants to endorse, except for the Compulsive scale (mean theta = 0.47). The scales with the lowest mean theta values (-1.88, -1.87, and -1.81) were Negativistic, Schizoid, and Sadistic, respectively, indicating that these scales were the least endorsed by participants. Although the average latent trait level of the participants was low, all scales showed adequate variability, suggesting that the sample comprised both people with healthy personalities and people with more pathological personality functioning.

Infit and outfit values express the correspondence between observed and expected theta values for the respondents assessed. The former is more sensitive to unexpected patterns of observations by people with theta values that are close to the item's difficulty level, and the latter is more sensitive to unexpected observations by people with theta values that are far from the item's

Table 1 – Descriptive statistics of the latent trait level of respondents and model fit indices (infit and outfit) (n = 350)

	Number of answered items	Theta	SE	Infit	Outfit
Depressive					
Mean (SD)	8.00 (0.30)	-1.52 (1.30)	0.62 (0.23)	0.99 (0.51)	1.00 (0.61)
Maximum	8.00	3.00	1.03	3.15	4.40
Minimum	4.00	-3.22	0.37	0.24	0.26
Schizoid					
Mean (SD)	8.00 (0.40)	-1.87 (1.20)	0.68 (0.20)	0.99 (0.70)	1.01 (0.90)
Maximum	8.00	3.23	1.06	4.22	6.22
Minimum	2.00	-3.54	0.44	0.10	0.13
Borderline					
Mean (SD)	7.00 (0.30)	-0.08 (1.25)	0.56 (0.09)	1.00 (0.68)	1.01 (0.75)
Maximum	7.00	3.60	1.07	4.41	5.61
Minimum	2.00	-3.69	0.51	0.09	0.12
Paranoid					
Mean (SD)	6.00 (0.10)	-1.43 (1.42)	0.73 (0.17)	0.95 (0.80)	1.00 (1.06)
Maximum	6.00	3.53	1.17	4.50	9.90
Minimum	4.00	-3.95	0.56	0.12	0.09
Sadistic					
Mean (SD)	6.00 (0.30)	-1.81 (1.17)	0.75 (0.19)	0.97 (0.73)	0.93 (0.77)
Maximum	6.00	3.13	1.08	4.68	5.35
Minimum	3.00	-3.36	0.52	0.18	0.17
Compulsive					
Mean (SD)	5.00 (0.20)	0.47 (1.39)	0.70 (0.11)	1.00 (0.75)	1.01 (0.77)
Maximum	5.00	3.42	1.09	4.23	4.58
Minimum	2.00	-3.50	0.63	0.08	0.08
Masochistic					
Mean (SD)	7.00 (0.30)	-1.60 (1.21)	0.67 (0.20)	0.97 (0.57)	0.96 (0.67)
Maximum	7.00	1.96	1.05	3.85	5.07
Minimum	3.00	-3.29	0.45	0.19	0.18
Antisocial					
Mean (SD)	7.00 (0.20)	-1.23 (1.05)	0.61 (0.20)	1.01 (0.57)	0.98 (0.63)
Maximum	7.00	2.06	1.05	3.73	5.76
Minimum	4.00	-2.99	0.42	0.09	0.09
Dependent					
Mean (SD)	6.00 (0.20)	-1.55 (1.20)	0.73 (0.21)	0.98 (0.71)	0.97 (0.80)
Maximum	6.00	1.70	1.14	4.33	5.43
Minimum	3.00	-3.21	0.48	0.13	0.13
Histrionic					
Mean (SD)	5.00 (0.20)	-1.21 (1.08)	0.72 (0.17)	0.99 (0.74)	1.00 (0.85)
Maximum	5.00	2.69	1.08	3.87	5.41
Minimum	2.00	-2.92	0.53	0.04	0.04
Negativistic					
Mean (SD)	6.00 (0.20)	-1.88 (1.07)	0.75 (0.19)	0.97 (0.63)	0.95 (0.78)
Maximum	6.00	1.50	1.13	3.68	5.28
Minimum	3.00	-3.31	0.52	0.23	0.24
Schizotypal					
Mean (SD)	5.00 (0.20)	-0.95 (1.20)	0.69 (0.18)	0.97 (0.72)	0.98 (0.78)
Maximum	5.00	2.63	1.08	3.89	4.39
Minimum	2.00	-2.78	0.53	0.02	0.02
Avoidant					
Mean (SD)	7.00 (0.30)	-1.38 (1.27)	0.64 (0.16)	0.98 (0.65)	1.01 (0.94)
Maximum	7.00	2.14	1.16	3.49	9.46
Minimum	2.00	-3.54	0.49	0.04	0.07
Narcissistic					
Mean (SD)	8.00 (0.30)	-0.55 (1.06)	0.51 (0.10)	1.01 (0.68)	1.00 (0.71)
Maximum	8.00	3.43	1.05	4.28	5.57
Minimum	3.00	-3.59	0.45	0.09	0.09

SD = standard deviation; SE = standard error.

difficulty level (e.g., when people with high ability miss easy items or people with low ability get hard items). These values in our sample were adequate because our average was below 1.20.[35] However, our maximum infit and outfit values exceeded 1.20, indicating that the expected model parameters could not adequately represent the responses observed in the scales for some subjects. Furthermore, the reliability of theta estimates calculated by the Rasch model ranged between 0.48 and 0.79, and 11 scales showed scores equal to or higher

than 0.60. These rates can be accounted as satisfactory, especially taking into consideration: 1) that each scale comprises only a small number of items, and 2) that the difficulty indices associated with endorsing each category are displaced in comparison to the subjects' theta levels, potentially influencing reliability parameters.[42]

Table 2 presents summarized descriptive statistics about the number of subjects who answered each item, the difficulty levels identified, infit and outfit indices, and item-theta correlations. The difficulty index associated with endorsing each item category, i.e., the threshold values between the scalar categories, varied between

Table 2 – Descriptive statistics of PDDI items, indices of difficulty, and adjustment

	n	b	SE	Infit	Outfit	Corr
Depressive						
Mean (SD)	348.40 (0.80)	0.00 (0.46)	0.10 (0.01)	1.05 (0.26)	1.05 (0.26)	
Maximum	350.00	1.02	0.12	1.39	1.39	
Minimum	348.00	-0.72	0.08	0.73	0.73	0.57-0.77
Schizoid						
Mean (SD)	348.1 (1.10)	0.00 (0.83)	0.10 (0.02)	1.09 (0.16)	1.01 (0.13)	
Maximum	350.0	1.29	0.14	1.29	1.30	
Minimum	346.0	-1.57	0.07	0.85	0.86	0.60-0.69
Borderline						
Mean (SD)	348.0 (0.90)	0.00 (0.85)	0.08 (0.00)	1.00 (0.09)	1.01 (0.11)	0.58-0.73
Maximum	350.0	1.27	0.08	1.17	1.21	
Minimum	347.0	-1.23	0.07	0.84	0.87	
Paranoid						
Mean (SD)	348.80 (0.90)	0.00 (1.21)	0.09 (0.01)	1.04 (0.22)	1.01 (0.28)	
Maximum	350.00	1.18	0.11	1.31	1.47	
Minimum	348.00	-2.54	0.08	0.70	0.67	0.56-0.76
Sadistic						
Mean (SD)	348.80 (0.90)	0.00 (0.81)	0.10 (0.02)	1.09 (0.15)	0.93 (0.14)	
Maximum	350.00	0.94	0.13	1.35	1.13	
Minimum	348.00	-1.15	0.08	0.93	0.73	0.59-0.72
Compulsive						
Mean (SD)	348.8 (1.00)	0.00 (0.42)	0.08 (0.00)	0.99 (0.09)	1.01 (0.08)	
Maximum	350.00	0.71	0.09	1.15	1.15	
Minimum	348.00	-0.51	0.08	0.91	0.92	0.66-0.75
Masochistic						
Mean (SD)	348.30 (0.70)	0.00 (0.79)	0.10 (0.02)	1.06 (0.20)	0.96 (0.11)	
Maximum	350.00	1.69	0.15	1.48	1.10	
Minimum	348.00	-0.88	0.08	0.83	0.80	0.40-0.74
Antisocial						
Mean (SD)	348.10 (0.40)	0.00 (0.45)	0.08 (0.01)	1.03 (0.17)	0.99 (0.22)	
Maximum	349.00	0.63	0.09	1.24	1.22	
Minimum	348.00	-0.76	0.07	0.78	0.62	0.50-0.68
Dependent						
Mean (SD)	348.70 (0.90)	0.00 (0.54)	0.10 (0.01)	1.05 (0.13)	0.97 (0.11)	
Maximum	350.0	0.62	0.11	1.23	1.13	
Minimum	348.0	-1.00	0.08	0.83	0.82	0.58-0.68
Histrionic						
Mean (SD)	348.20 (0.40)	0.00 (0.53)	0.09 (0.01)	1.03 (0.18)	1.00 (0.13)	
Maximum	349.00	0.94	0.10	1.38	1.23	
Minimum	348.00	-0.57	0.08	0.87	0.86	0.54-0.72
Negativistic						
Mean (SD)	348.00 (0.60)	0.00 (0.87)	0.11 (0.02)	1.07 (0.13)	0.95 (0.15)	
Maximum	349.00	1.41	0.15	1.18	1.23	
Minimum	347.00	-1.08	0.09	0.79	0.78	0.48-0.74
Schizotypal						
Mean (SD)	347.04 (1.70)	0.00 (0.57)	0.08 (0.01)	1.02 (0.31)	0.98 (0.29)	
Maximum	349.00	0.67	0.09	1.62	1.53	
Minimum	344.00	-0.76	0.07	0.81	0.76	0.58-0.75
Avoidant						
Mean (SD)	348.60 (0.90)	0.00 (0.97)	0.09 (0.01)	1.07 (0.20)	1.01 (0.16)	
Maximum	350.00	1.34	0.12	1.37	1.28	
Minimum	348.00	-1.48	0.07	0.80	0.84	0.55-0.72
Narcissistic						
Mean (SD)	348.60 (0.90)	0.00 (0.82)	0.07 (0.00)	1.00 (0.09)	1.00 (0.12)	
Maximum	350.00	0.90	0.08	1.23	1.30	
Minimum	348.00	-1.58	0.07	0.93	0.87	0.53-0.65

b = item difficulty; Corr = item-total correlations; n = number of respondents; SD = standard deviation; SE = standard error.

-2.54 (minimum b for Paranoid) and 1.69 (maximum b for Masochistic), with standard deviations around one logit (i.e., 1.0). These findings suggest that PDDI items cover a wide range of the construct (personality disorders). Average fit indices obtained for the items were adequate, i.e., below 1.20. However, our maximum infit and outfit indices suggest that at least one item in each scale presented incongruity between what was expected by the model and the observed data, with the exception of the Compulsive scale. Moreover, item-total correlations pointed to high magnitudes of correlation between groups of items and their scales. Table 2 presents only the low and high correlations obtained for each scale. Figure 2 provides illustrative data on the analysis of response categories in the Paranoid scale.

Figure 2 shows the probabilities of the respondents' endorsing each response category and how these are mapped at different levels of theta values to an item with
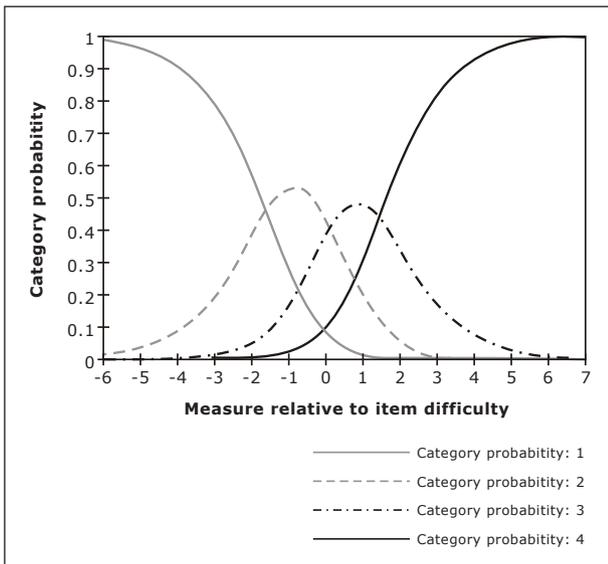
$b_i = 0$ (i.e., difficulty level equal to zero). Intersections between two categories can be understood as the threshold value of the transition between the categories. The threshold between categories 1 and 2 was equal to -1.58, between 2 and 3, 0.17, and between 3 and 4, 1.41. A clear representation of all categories was observed. Non-overlapping curve peaks, i.e., the separation of the different regions associated with each scalar value, are a desirable metric feature; these empirical data show that responses to the items could be modeled quantitatively by means of a monotonic increasing relationship between theta values and scalar categories.

Figure 3, the construct map, illustrates one of the most important applications of the Rasch model in the evaluation of disorders, one that can help overcome the arbitrary metrics problem. The item map analysis exemplified in Figure 3 for the Paranoid scale shows the responses expected for each item according to different trait levels (theta values, plotted as the top and bottom lines of the box, ranging from -6 to +4). Items are displayed in the map (inside the boxes on the right) starting with those that are easier for participants to endorse (bottom, item no. 13) up to those that are harder to endorse (top, item no. 26). Item number and description are displayed on the right margin. In the body of the map, response categories are shown ranging from 1 ("does not describe me") to 4 ("describes me extremely or completely") for each item; their positions are given by the thresholds in the theta range.

At the bottom of the map, the distribution of respondents is presented (the number of respondents in each band should be read vertically, e.g., 27 at theta = -4), as well as sample percentiles. Just below the distribution of participants, the letters T, S, and M appear, referring to the mean (M), one standard deviation above or below the mean (S), and two standard deviations above or below the mean (T). For the scope of the present study and given the theoretical perspective in

Figure 2 – Category probability curves of the Paranoid scale
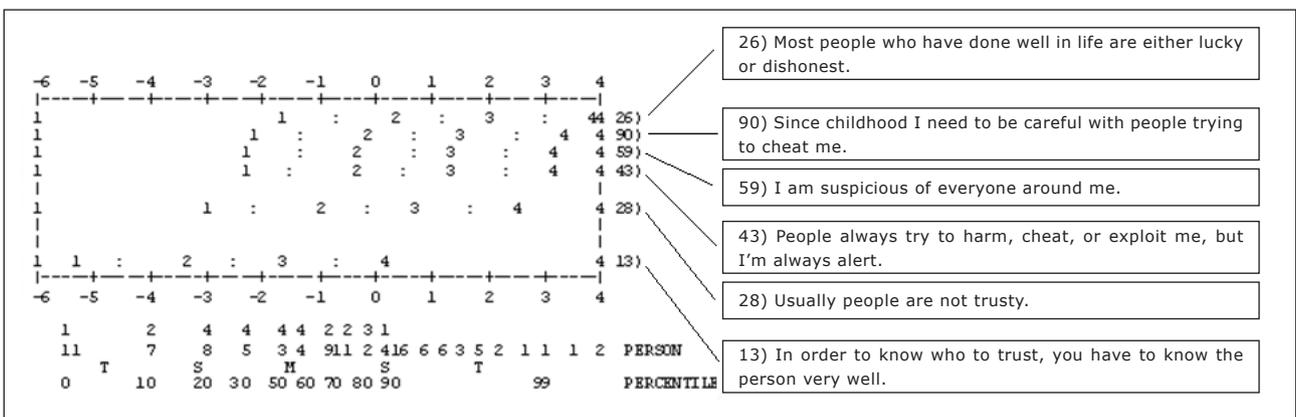
Figure 3 – Person-item map in the Paranoid scale

relation to the construct being measured, a qualitative analysis was completed for all Paranoid scale items, in an attempt to demonstrate the clinical contribution of these items.

A higher concentration of respondents with theta values ranging from -4.0 to 0 can be observed, which was an expected finding taking into consideration the average theta in this sample, of -1.43 (Table 1). In general, the content of the items in the Paranoid scale refer, more or less directly, to the respondents' difficulty in trusting others, which appears to be a central feature in paranoid personality disorder.[3] The hierarchical arrangement of items suggests that items no. 13 and 28, whose contents do not contain absolute statements ("enough" and "usually"), tended to be more easily endorsed by participants. The next two items (no. 43 and 59) include absolute statements ("always" and "all"), indicating greater rigidity when compared with the previous items. Item no. 90 refers to childhood; respondents endorsing this item believed they had this personality characteristic for as long as they could remember. Finally, item no. 26 is related to a dimension of paranoid functioning, namely, discrediting the abilities of others[3] and difficulty believing that people can be successful without harming others.

Thus, according to the analyses here described, people with certain latent trait levels (i.e., paranoid functioning levels) tend to agree with the assertions contained in the corresponding scale. For example, people with a theta equal to zero tend to agree only with the first two items, while people with a theta equal to 0.5 tend to disagree only with item 26. This 0.5 difference between the two theta values, used as examples, points to substantial changes in the functioning of these people. Accordingly, the standardized scalar number (theta) is not just an arbitrary number on the scale; rather, it allows to infer which characteristics subjects at a given latent trait level potentially have and which they are unlikely to have.[24,35]

Conclusions

This study aimed to describe item and person parameters obtained with the Rasch model for an instrument designed to assess personality disorders based on Millon's theory. Overall, analysis with the Rasch model revealed that the PDDI has adequate psychometric properties for the assessment of personality disorders. Among the contributions of IRT for clinical instruments, the person-item map deserves to be highlighted as a successful attempt to improve the understanding of clinical scores obtained in response to particular test items.

Among the limitations of this study, two are especially important: 1) considering that the PDDI was developed to measure pathological personality characteristics, the relatively small number of confirmed psychiatric cases may have affected the psychometric data of the test; and the small number of items on some scales, such as the Paranoid scale, which could be increased in an attempt to cover other aspects of functioning typical of personality disorders. Notwithstanding, we hope that our study will contribute relevant data to the field of personality disorder assessment in Brazil, especially from the perspective of modern psychometrics, which is already widely used in other countries in the same field of study.

References

1. Millon T, Millon CM, Meagher S, Grossman S, Ramanath R. Personality disorders in modern life. New Jersey: Wiley; 2004.
2. Millon T, Grossman S. Overcoming resistant personality disorders. New Jersey: Wiley; 2007.
3. Millon T, Grossman S. Moderating severe personality disorders. New Jersey: Wiley; 2007.
4. Millon T, Grossman S, Tringone R. The Millon personality spectrometer: a tool for personality spectrum analyses, diagnoses, and treatments. In: Millon T, Krueger RF, Simonsen E, editors. Contemporary directions in psychopathology: scientific foundations of the DSM-V and ICD-11. New York: Guilford; 2010. p. 391-418.
5. Millon T, Davis RD. Disorders of personality: DSM-IV and beyond. New Jersey: Wiley; 1996.
6. Davis RD. Millon: essentials of his science, theory, classification, assessment, and therapy. J Pers Assess. 1999;72:330-52.
7. Alchieri JC. Modelo dos estilos de personalidade de Millon: adaptação do Inventário Millon de Estilos de Personalidade [tese]. Porto Alegre: Universidade Federal do Rio Grande do Sul; 2004.
8. Strack S, Millon T. Contributions to the dimensional assessment of personality disorders using Millon's model and the Millon Clinical Multiaxial Inventory (MCMI9-III). J Pers Assess. 2007;89:56-69.
9. Millon T, Millon CM, Davis RD. MCMI-III manual. Minneapolis: Dicandrien; 1994.
10. Craig RJ, Bivens A. Factor structure of the MCMI-III. J Pers Assess. 1998;70:190-6.
11. Dyce JA, O'Connor BP, Parkins S, Janzen H. Correlational structure of the MCMI-III personality disorder scales and comparison with other data sets. J Pers Assess. 1997;69:568-82.
12. Rossi G, Van der Ark LA, Sloore H. Factor analysis of the Dutch-language version of the MCMI-III. J Pers Assess. 2007;88:144-57.
13. Rossi G, Brande IV, Tobac A, Sloore H, Hauben C. Convergent validity of the MCMI-III personality disorder scales and the MMPI-2 scales. J Pers Disord. 2003;17:330-40.
14. Handler L, Meyer GJ. The importance of teaching and learning personality assessment. In: Handler L, Hilsenroth M, editors. Teaching and learning personality assessment. New Jersey: Lawrence Erlbaum; 1997. p. 3-30.
15. Widiger TA, Trull TJ. Place tectonics in the classification of personality disorder: shifting to a dimensional model. Am Psychol. 2007;62:71-83.

16. Morana HCP. Identificação do ponto de corte para a escala PCL-R (Psychopathy Checklist Revised) em população forense brasileira: caracterização de dois subtipos da personalidade; transtorno global e parcial [tese]. São Paulo: Universidade de São Paulo; 2003.

17. Carvalho LF. Construção e validação do Inventário Dimensional dos Transtornos da Personalidade [dissertação]. Itatiba: Universidade São Francisco; 2008.

18. Carvalho LF, Bartholomeu D, Silva MCR. Instrumentos para avaliação dos transtornos da personalidade no Brasil. Aval Psicol. 2010;9:289-98.

19. Samuel D, Simms LJ, Clark LA, Livesley J, Widiger TA. An item response theory integration of normal and abnormal personality scales. Pers Disord Theory Res Treat. 2010;1:5-21.

20. Walton KE, Roberts BW, Krueger RF, Blonigen DM, Hicks BM. Capturing abnormal personality with normal personality inventories: an item response theory approach. J Pers. 2008;76:1623-47.

21. Primi R. Avanços na interpretação de escalas com a aplicação da teoria de resposta ao item. Aval Psicol. 2004;3:53-8.

22. Pasquali L, Primi R. Fundamentos da teoria da resposta ao item: TRI. Aval Psicol. 2003;2:99-110.

23. Pasquali L, Primi R. Fundamentos da teoria de resposta ao item – TRI. In: Pasquali L. Teoria de resposta ao item: teoria, procedimentos e aplicações. Brasília: LabPAM/Unb; 2007. p. 11-52.

24. Embretson SE, Reise SP. Item response theory for psychologists. Mahwah: Lawrence Erlbaum; 2000.

25. Wright BD, Masters GN. Rating scale analysis. Chicago: MESA; 1982.

26. Andrich D. A rating formulation for ordered response categories. Psychometrika. 1978;43:561-73.

27. Linacre JM. The partial credit model and the one-item rating scale model. Rasch Meas Trans. 2005;19:1000-2.

28. Nunes CHSS, Primi R, Nunes MFO, Muniz M, Cunha TF, Couto G. Teoria de resposta ao item para otimização de escalas tipo Likert: um exemplo de aplicação. Rev Iberoam Diagn Eval Psicol. 2008;1:51-79.

29. Balsis S, Gleason MEJ, Woods CM, Oltmanns TF. Age group bias in DSM-IV personality disorder criteria: an item response theory analysis. Psychol Aging. 2007;22:171-85.

30. Cooke DJ, Michie C. An item response theory analysis of the Hare Psychopathy Checklist--Revised. Psychol Assess. 1997;9:3-14.

31. Feske U, Kirisci L, Tarter RE, Pilkonis PA. An application of item response theory to the DSM-III-R criteria for borderline personality disorder. J Pers Disord. 2007;21:418-33.

32. Olatunji BO, Woods C, Jong PJ, Teachman B, Sawchuk, CN, David B. Development and initial validation of an abbreviated Spider Phobia Questionnaire using item response theory. Behav Ther. 2009;40:114-30.

33. Stelmack J, Szlyk JP, Stelmack T, Babcock-Parziale J, Demers-Turco P, Williams TR, et al. Use of Rasch person-item map in exploratory data analysis: a clinical perspective. J Rehabil Res Dev. 2004;41:233-41.

34. Blanton H, Jaccard J. Arbitrary metrics in psychology. Am Psychol. 2006;61:27-41.

35. Embretson SE. The continued search for nonarbitrary metrics in psychology. Am Psychol. 2006;61:50-5.

36. Primi R, Carvalho LF, Miguel FK, Muniz M. Resultado dos fatores da BFP por meio da teoria de resposta ao item: interpretação referenciada no item. In: Nunes CH, Hutz C, Nunes MF, organizadores. Bateria Fatorial de Personalidade (BFP): manual técnico. São Paulo: Casa do Psicólogo; 2010. p. 153-70.

37. Reise SP, Haviland MG. Item response theory and the measurement of clinical change. J Pers Assess. 2005;84:228-38.

38. Smith Jr EV. Introduction to Rasch measurement theory, models and applications. In: Smith Jr EV, Smith RM. Introduction to Rasch measurement. Maple Glover: JAM; 2004. p. 1-24.

39. Carvalho LF, Primi R. Development of a Brazilian inventory for the assessment of personality disorders based on Millon's model. Society for Personality Assessment Annual Meeting; 2010; California, USA.

40. Carvalho LF, Primi R. Personality style assessment in patients with chronic pain. Society for Personality Assessment Annual Meeting; 2009; Chicago, USA.

41. Winsteps Rasch Measurement. Winsteps version 3.69.1. Chicago: Winsteps; 2009.

42. Linacre JM, Wright BD. Reasonable mean-square fit values. Rasch Meas Trans. 1994;8:370.

Correspondence
Lucas de Francisco Carvalho
Universidade São Francisco
Rua Alexandre Rodrigues Barbosa, 45
13251-900 - Itatiba, SP - Brazil
E-mail: lucas@labape.com.br