Revista Colombiana de Psicología
ISSN: 0121-5469
revpsico_fchbog@unal.edu.co
Universidad Nacional de Colombia
Colombia

CARVALHO, LUCAS DE FRANCISCO; QUEIROGA DE OLIVEIRA FILHO, ANANIAS; PESSOTTO, FERNANDO; VINCENZI BORTOLOTTI, SILVANA LIGIA

Application of the Unfolding Model to the Aggression Dimension of the Dimensional Clinical Personality Inventory (IDCP)

Revista Colombiana de Psicología, vol. 23, núm. 2, julio-diciembre, 2014, pp. 339-349
Universidad Nacional de Colombia
Bogotá, Colombia

Available in: http://www.redalyc.org/articulo.oa?id=80434236008

doi: 10.15446/rcp.v23n2.41428

Application of the Unfolding Model to the Aggression Dimension of the Dimensional Clinical Personality Inventory (IDCP)*

Aplicación del Modelo de Desdoblamiento a la Dimensión de Agresividad del Inventario Dimensional Clínico de Personalidad (IDCP)

Aplicação do Modelo de Desdobramento à Dimensão Agressividade do Inventário Dimensional Clínico da Personalidade (IDCP)

LUCAS DE FRANCISCO CARVALHO
ANANIAS QUEIROGA DE OLIVEIRA FILHO
FERNANDO PESSOTTO
Universidade São Francisco, Itatiba, São Paulo, Brasil

SILVANA LIGIA VINCENZI BORTOLOTTI
Universidade Tecnológica Federal do Paraná, Medianeira, Paraná, Brasil

Abstract

This article compares the suitability of the dominance and unfolding models for the analysis of the Aggressiveness dimension in the IDCP (Dimensional Clinical Personality Inventory). The study included 975 subjects, with ages ranging from 18 to 81 years (M=29.82, SD=12.28), 58.9% of which were women. The IDCP is composed of 163 items and 12 dimensions; 27 items are related to Aggression. The analysis with the unfolding model indicated the exclusion of 15 items due to standard error. Results showed a better fit for the dominance model. This result may be due to the nature of the construct, because the items assess pathological aspects of personality representing one end of the continuum.

Keywords: dominance model, unidimensional unfolding model, personality, personality disorders, item-response theory.

Resumen

Este estudio compara la idoneidad de los modelos de dominancia y desdoblamiento para el análisis de la dimensión de Agresividad del IDCP (Inventario Dimensional Clínico de Personalidad). Participaron en el estudio 975 sujetos entre los 18 y los 81 años de edad (M=29,82; DE=12,28), de los cuales el 58,9% eran mujeres. El IDCP está integrado por 163 ítems y 12 dimensiones, con 27 ítems referentes a la Agresividad. El análisis a través del modelo del desdoblamiento produjo la exclusión de 15 ítems debido a error estándar. Los resultados mostraron un mejor ajuste del modelo de dominancia. Este resultado puede deberse a la naturaleza del constructo, por cuanto los ítems evalúan los aspectos patológicos de la personalidad que representan un extremo del continuum.

Palabras clave: modelo de dominancia, modelo de desdoblamiento, personalidad, trastornos de la personalidad, teoría de respuesta a un ítem.

Resumo

Este artigo compara a adequação da dominância e dos modelos que se desdobram para a análise da dimensão da agressividade no IDCP (Inventário Dimensional Clínico da Personalidade). O estudo incluiu 975 indivíduos, com idades que variam dos 18 até os 81 anos (M=29,82; SD=12,28), 58,9% dos quais eram mulheres. O IDCP é composto de 163 itens e 12 dimensões; 27 itens estão relacionados com a agressão. A análise com o modelo de desdobramento indicou a exclusão de 15 itens devido ao padrão de erro. Os resultados mostraram um melhor ajuste para o modelo de dominação. Esse resultado pode ser devido à natureza da construção, porque os itens avaliaram aspectos patológicos da personalidade, o que representa uma extremidade do continuum.

Palavras-chave: modelo da dominância, modelo de desdobramento unidimensional, personalidade, transtornos de personalidade, teoria da resposta ao item.

How to cite this article: Carvalho, L. de F., Filho, A. Q. de O., Pessoto, F., & Bortolotti, S. L. V. (2014). Application of the unfolding model to the aggression dimension of the Dimensional Clinical Personality Inventory (IDCP). *Revista Colombiana de Psicología, 23*(2), 339-349. doi: 10.15446/rcp.v23n2.41428.

Correspondence concerning this article should be adressed to Lucas de Francisco Carvalho, e-mail: lucas@labape.com.br. Universidade São Francisco, Rua Alexandre Rodrigues Barbosa, 45 CEP 13251-900, Itatiba, São Paulo, Brasil.

SCIENTIFIC RESEARCH ARTICLE
RECEIVED: 7 JANUARY 2014 - ACCEPTED: 24 AUGUST 2014

* This research project was funded by CNPq and FAPESP.

IN THE Item Response Theory (IRT) and in psychometrics in general, it is assumed that items that operationalize the constructs function according to the principles of the dominance model paradigm. In such paradigm, the more present and/or intense the construct determined in the individual is, the greater the likelihood that he or she will agree or give the correct answers for the items representing that construct. However, there is empirical evidence (e.g., Roberts, Donoghue, & Laughlin, 2000) suggesting that not all constructs or items adequately fit the paradigm of dominance, indicating the need to use other instruments, such as the unfolding model.

In the unfolding model it is assumed that the determination of the probability of the individuals' response to an item is more related to the proximity of the location of the item and the subject, one in relation to the other, and not necessarily to the intensity of the representation of the construct. Accordingly, the probability of agreeing with an item increases as the level in the latent construct of the individual approaches the level on the item's construct. The present study verifies the postulation of the unfolding model, which fits into the family of mathematical models of IRT, investigating its suitability as compared to the dominance of a set of items assessing personality characteristics of the model.

The IRT has been widely used in the development of instruments. With regard to studies of personality, using the dominance model is a common strategy (Meijer & Baneke, 2004). In dominance models it is assumed that the higher the score on the latent trait, the higher the probability of getting the correct response in the case of skills, or of agreeing with the item, in the case of non-cognitive constructs. It is noteworthy that the present study focused on non-cognitive constructs. For instance, when measuring depression using items such as "I am often depressed" within the paradigm of dominance, it is reasonable to assume that the more depressed someone is, the greater the likelihood that he/she will agree with the item.

In models of dominance, the information function of the scale indicates whether the measurement accuracy of the latent trait is higher or lower, because it is inversely related to the standard error of measurement (Embretson & Reise, 2000). The models of dominance are usually based on the use of one, two, or three parameters to assess the skill level (θ) of respondents. The model with one parameter, or Rasch model, makes use of the difficulty (b) index, which is related to the probability of hit/agreement with the item, representing the level of the latent trait necessary to properly respond to it.

In the two-parameter model, besides the difficulty index, the discrimination power item (a), which allows differentiating individuals based on knowledge of a particular subject or skill at some task, is also part of the equation. In the one-parameter model, it is assumed that the power of discrimination is the same for all items. Finally, the third model includes a parameter that represents the probability of hit by chance (c), indicating an increase in the probability of a subject —with a certain level of the latent trait— choosing a given response, attributing this increase randomly. In all dominance models the Item Characteristic Curves (ICC) are described by a monotonically increasing function.

Unlike the paradigm of dominance, a model less usually found in the literature is the so-called unfolding or ideal point model. Conceptually, the unfolding model implies that a subject must agree with a certain item if both are closely located in the continuum of the latent construct. Therefore, in the unfolding model the ICC is not represented by a monotonically increasing function; the curve is determined by the proximity of the item and the respondent to the level of latent construct, rather than by the increased likelihood of agreement according to the intensity presented by the individual in this construct (Coombs, 1964; Roberts, Donoghue, & Laughlin, 1999).

$$P(Z = z|\Theta) = \frac{exp\,[\alpha_i\,(\,z(\Theta_j - \delta_i) - \Sigma_{k=0}^{z}\tau_{ik}\,)] + exp\,[\alpha_i((M\text{-}Z)\,(\Theta_j - \delta_i) - \Sigma_{k=0}^{z}\tau_{ik}\,)]}{\Sigma_{V=0}^{H}\,[exp\,(\alpha_i\,[v(\Theta_j - \delta_i) - \Sigma_{k=0}^{v}\tau_{ik}\,]) + exp(\alpha_i\,[(M\text{-}Z)\,(\Theta_j - \delta_i) - \Sigma_{k=0}^{z}\tau_{ik}\,])]}$$

Figure 1. Item Characteristic Curves equation according to the Generalized Graded Unfolding Model

Z being the probability of an observable subjective response;
i representing the item given the possible options;
z, options of objective responses;
θ_i, skill level;
j, position of the person in relation to θ;
δ_i, difficulty of the item;
α_i, discrimination power of the item;
τ_{ik}, threshold category of the *K*-th option (where *K* is a number minus the number of possible response options) or the level of the response.

The Generalized Graded Unfolding Model (GGUM) was developed by Roberts et al. in 2000. Although other types of developments are being used with relative success, the GGUM has received considerably more attention in the literature and was chosen for analysis in this research model. Mathematically, the equation of ICC on the GGUM model is represented as shown in Figure 1.

Although the equation seems similar to the typical models of IRT, the interpretation is quite different. The end, or location of a point, is represented by considering the attribute level at which the item "unfolds" and begins to show a non-monotonic behavior. In this model, people with an attribute level equal to δ are most likely to corroborate the item, and this probability decreases for those with θ levels below or above δ. Roberts, Lin and Laughlin (2001) explain that a person cannot corroborate an item if it presents an excessively negative content, being its δ parameter much smaller than the parameter θ of the individual. The same may occur in the event of an excessively positive content, since the δ parameter of the item is considerably higher than the parameter θ of the individual. Thus, level θ is so different in relation to level δ that the subject ends up not corroborating the item. That is, the individual may disagree by lack or excess of θ; in these two situations the responses are called subjective responses.

In other words, the τ parameters are distributed symmetrically around the point (θ – δ = 0). Thus, to parameterize the GGUM it is assumed that for all options of objective responses (for instance, "I disagree"), there are two options of subjective responses: subjects may disagree by being far above the position of the item or far below it). On the scale of the latent trait, "above" would be "to the right" of the item and "below" "to its left". Consider the item to assess the resistance to organizational change (Bortolotti, 2010): "When changes happen, try to do only what is necessary". This item can be disagreed by individuals who have no resistance to change (disagree above the item). For instance, those who accept changes, or by individuals that are resistant to change (disagree below the item). Likewise, the number of thresholds and the possibility of subjective responses (subjective response curves), M, is equal to 2 (H+1), where H is the most extreme option. It is assumed that these values are symmetrical around the point (δ – θ = 0), being half negative and half positive in terms of value. Operationally, only half of the negative values are used in the calculation of probabilities of response because their positive valued reciprocals are redundant.

We observed, then, that the dominance proposals typically used in psychology in analyses based on the IRT, and the unfolding proposals present clearly distinct postulations. Despite

the still limited literature in the area, there are some studies using the unfolding model for the verification of adequacy of items assessing non-cognitive constructs, specifically personality (Chernyshenko, Stark, Chan, Drasgow, & Williams, 2001; Conn & Rieke, 1994; Meijer & Baneke, 2004; Stark, Chernyshenko, Drasgow, & Williams, 2006). These studies demonstrate the applicability of the model and also that some items or sets of items are more appropriate for the unfolding proposal than for the dominance model (Wang, Tay, & Drasgow, 2013). Basically, it is agreed that the better adequacy of these items depends on the nature of their content. In general, items that tend to have neutral content (and not clearly directed to an extreme) tend to fit more appropriately the unfolding models. However, specific, more adequate constructs for one given model are not reported in the literature.

The objective of the present study was to verify the adequacy of a set of items in assessing a dimension of personality, Aggressiveness (Dimensional Clinical Personality Inventory - IDCP), using a model based on the dominance paradigm (graduated response model from the Rasch model family – andrich rating scale model) and a model based on the unfolding paradigm (GGUM). Furthermore, we sought to compare the levels of adequacy and discuss the possibilities of use of the unfolding paradigm, which is seldom used in psychology in Brazil. We specifically chose this set of items, because: (a) it is necessary for the set of items to be unidimensional; (b) this dimension is composed of a number of items sufficient for the analysis; other dimensions of the IDCP present a very restricted number of items and we aimed at verifying the possibility of using this model in a specific set of items for an specific construct, focused more on the model than on the instrument itself.

Method

Participants

For the present study, 975 individuals were selected by convenience sampling, with ages ranging from 18 to 81 years (M=29.82; SD=12.28), of which 574 were women (58.9%). Most participants were university students (63.2%), followed by high school students (11.4%), and 168 subjects (17.2%) did not inform their level of education.

Instrument

The IDCP (Carvalho & Primi, 2011) was developed based on Millon's theory and on diagnosis characteristics from the DSM-IV-TR Axis II (APA - American Psychiatric Association, 2002) with the objective of analyzing pathological personality characteristics in adults (Carvalho, 2011). The instrument is composed of 163 items distributed in 12 dimensions: Dependence (inability to make decisions with inadequate self-performance beliefs), Aggressiveness (disregard to others and violent acts), Mood Instability (mood oscillation with irritability, sadness, and guilt), Eccentricity (eccentric behaviors, belief of being different and not taking pleasure in being with other people), Need for Attention (search for friendships, seduction and exaggerated need for attention from others), Distrust (inability to trust people and ideas of persecution), Grandiosity (exaggerated need for recognition and admiration with exaggerated beliefs in one's own merit and superiority), Isolation (preference for being alone with little pleasure in relationships), Avoidance of Criticisms (beliefs of inability, humiliation, and criticism by others), Self-Sacrifice (self-disrespect, helping despite harms to himself), Conscientiousness (need for organization, order, and perfectionism and excessive concern), and Impulsivity (inconsequence, breaking laws, and engaging in dangerous activities).

Items must be answered using a 4-point Likert-like scale, ranging from *does not describe me at all* (1) to *it describes me accurately* (4). Studies were conducted to search for evidence of validity and reliability and, according to Carvalho (2011), the results obtained tended to be adequate. The approximate time of application is 20 min, and the instrument can be collectively or individually administered.

Procedure

This research was initially submitted to a Research Ethics Committee and subsequently approved (C.A.A.E: 0144.0.142.000-07). After approval, participants were recruited in classrooms or on the campus of private universities in the city of São Paulo. Individuals were invited to participate in the research by means of an Informed Consent Term, and participants were only included in the study after having read and signed the document. During data collection, at least one of the authors was present to clarify possible questions.

Data Analysis

After collection, data were tabulated and analyzed statistically. Considering the objectives of this study, we proceeded with the implementation of the graduated response (dominance model), using the Winsteps software with the mean difficulty set to zero, and the GGUM (unfolding model), which was done with the GGUM 2004 software. To run the data with the GGUM software, a quadrature equal to 50 was set. The steps for calibration are presented in the sequence.

The following aspects were verified: the dimensionality of the set of items; the operation of response categories; the summarized descriptive statistics of the latent traits (theta) of the respondents, their respective fit indices (infit and outfit) and data regarding the items; the level of difficulty; the item-theta correlation; reliability indices. For both models, we sought to obtain unidimensionality and an adequate operation of the response categories. Specifically regarding the unfolding model, we also used as a criterion a minimum number of items (15) and of the sample (750). In the case of the dominance model, we observed the fit indices and the item-theta correlation to verify the suitability of the model; in the case of the unfolding model, these data were also observed, but to be kept in the analysis, items had to show a similar level of standard error in relation to the other items of the set.

Results and Discussion

Given the objectives of the present study, we initially verified the unidimensionality assumption of the set of items through the analysis of the main components of residues implemented on the Winsteps. In the IRT the unidimensionality assumption implies that every non-random variation found be explained by a single dimension of difficulty and skill. The one-parameter model predicts the probability of success based on the difference between the ability of a person and the difficulty of the task. Unlikely responses are expected, with a low frequency of random occurrences. However, it is more likely that no strictly unidimensional structures be found in the psychological variables. This may be due to factors such as variability of content, complexity of constructs, and varied shapes of the items (Baker, 2001). Wright and Stone (2004) add that it is not possible to verify whether a test is strictly unidimensional. They propose, however, that it is possible to test whether or not other dimensions do not generate significant distortions.

The principal component analysis performed by Winsteps is conducted with this new matrix based on the proportion of responses that are not predicted by the model. From the parameters of the items and the subjects it is possible to calculate an expected response for each individual at each item. The discrepancy between the modeled (expected) response and

the actual response is considered as waste. Thus, if a component presents items with a magnitude greater than 2.0 (Linacre, 2009), it is suggested that a second dimension can potentially affect the data by obscuring the meaning of the first dimension. In the present study eigenvalues for waste equal to or greater than 2.0 were not found, indicating the unidimensionality of the set of items.

Once the unidimensionality was observed, we proceeded to verify the functioning of the response categories of the set of items. The analysis of response categories allowed us to determine whether the categories on the Likert-like scale proposed for the test obtained a minimally adequate functioning. Figure 2 provides illustrative data about the response categories of the Aggression dimension.

Figure 2 shows the response categories of the items; the x-axis shows the theta scale (level of respondents in latent trait) and the y-axis shows the probability of participant response at different levels of theta (in the figure the *b* average was set at zero). The figure also presents the probabilities of participant endorsement in each of the response categories, and their distributions at different levels of theta for an item *bi=o* (i.e., average level of difficulty equal to zero). The intersection between two categories can be interpreted as the threshold value between those categories. No overlapping of the curves was observed, suggesting that all response categories showed a region in the theta (horizontal axis) in which they would be the most likely.

Besides the visual analysis of the response categories, it is also important to investigate whether there is a clear progression in the theta scale on the passage between one response category and another. There was a clear progression of the levels of theta as the Likert scale increased, with a mean theta in the passage between categories 1-2 equal to -.30; 2-3 equal to .05; and 3-4 equal to .25.

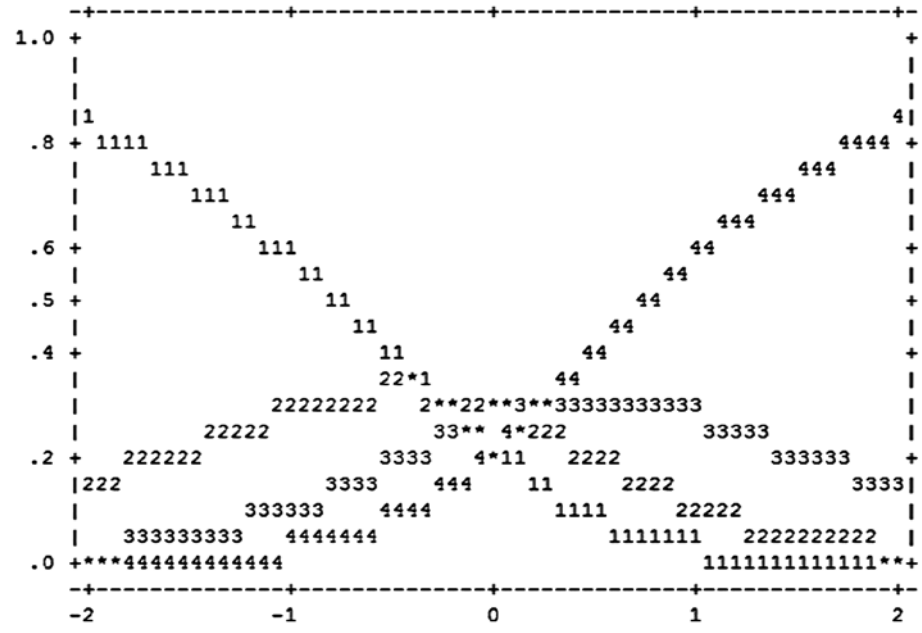

Figure 2. Response categories in the Aggression dimension.

For application of the models of graduated response and GGUM, the assumptions of unidimensionality and distribution of response categories were checked. Specifically for the GGUM, three further criteria were observed: whether the response categories had a frequency higher than zero; whether the set of items comprised at least 15 items; and whether the sample comprised at least 750 individuals (Roberts et al., 2000). In the present study all response categories had frequencies greater than zero, a set of 27 items was used, and, as already explained, the sample had an N greater than 750, so all the criteria were met.

After running the GGUM model, we assessed the standard error of the items, which was an exclusion criterion (Chernyshenko, Stark, Drasgow, & Roberts, 2007). In a first analysis we observed that two items obtained a discrepant standard error, that is, they were much higher in relation to the others (for this model there is no cutoff point described in the literature) and so, after the removal of the items, we proceeded to a new analysis. It is worth noticing that the exclusion of the items was conducted manually, and not automatically by the software. This indicated that 13 items had very different standard errors, with 15 items excluded up to this moment. Then, an analysis of the 12 remaining items was conducted, and the results were presented as follows:

In Table 1 we can observe summarized descriptive statistics of the latent traits (theta) of the respondents, their respective fit indices (infit and outfit), and the number of items responded to in the Aggressiveness dimension of the IDCP. Furthermore, this table summarizes the descriptive data for the items, that is, the level of difficulty, the fit indices, the item-theta correlation, and the reliability indexes (real and modulated).

In the case of the graduated response model, given the mean theta (negative) of the respondents, the mean of the items suggests that the subjects tended to disagree. The mean level of the latent trait in the graduated model of response suggests that the items are above the mean, indicating a greater presence of less endorsed items (*b* positive). The standard error of theta suggests that the sample is composed of people with distinct levels of intensity in relation to the characteristics assessed by the set of items. To make this inference we assumed that the lighter or more extreme scores observed are indicative of the level of functioning of the subjects regarding aggressiveness. In the GGUM, in contrast to the graduated response model, the mean of the latent trait indicates that if the items are close to the mean, they follow an approximate distribution of more endorsed and less endorsed items. Based on the theta standard error we may conclude, same as with the model of graduated response, that the participants have different levels of the latent trait.

Table 1
Descriptive summary statistics of persons and items

		Persons			Items			Theta correlation	Reliability
		Theta	Infit	Outfit	*b*	Infit	Outfit		
GGUM	*M (SD)*	0.32 (0.67)			-1.02 (0.56)	0.92 (0.07)	0.84 (0.19)		
	Max.	-0.75			-0.52	1	1	-.07 - 0.61	--
	Min.	1.52			-2.65	0.4	0.29		
Rating Scale	*M (SD)*	-1.39 (0.90)	1.03 (0.41)	0.98 (0.51)	0 (0.46)	1.05 (0.15)	0.98 (.22)		
	Max.	2	2.71	4.54	1.21	1.40	1.50	.32 - .56	.79 (0.81)
	Min.	-3.71	0.19	0.20	-0.76	0.77	0.67		

Note: Index for infit and outfit (*M*, *SD*, minimum and maximum), as well as the reliability, were not presented for the GGUM model because the software does not provide this data.

By using the infit and outfit indices in the model of graduated response we verified discrepancies between the expected values and those observed in relation to the estimation of the respondents' thetas. These values tended to be adequate (Linacre & Wright, 1994) once the mean was below 1.3. However, for some subjects, maximum values of fit indices higher than 1.3 were found, suggesting discrepancies beyond that expected by the model. The reliability index of the theta estimates calculated by the model of graduate response may be considered satisfactory (Embretson, 2000). In the GGUM model, the fit analysis was conducted based on the outfit and infit indices. The results obtained indicate a good fit of the items, as none of them was higher than 1.3. It should be highlighted that one item presented an outfit lower than 0.5, which indicated a better fit than that expected by the model. Although it may suggest a possible distortion of the model, this parameter was not used as an exclusion criterion for the items. This information can be observed in Table 2.

Table 2
Descriptive data of the items of in the Aggression dimension

	Dominance model				Unfolding model					
Item	*b*	**Infit**	**Outfit**	**Item - theta correlation**	*a*	**delta(δ)**	**Infit**	**Outfit**	**Item - theta correlation**	**Discrimination (α)**
A016	1.21	1.31	1.08	.32	1.05	-	-	-	-	-
B123	1.02	1.21	0.76	.39	1.10	-1.27	0.74	0.29	-.7	4.03
B114	0.60	1.09	0.87	.44	1.10	-1.12	0.92	0.87	.55	1.33
B159	0.45	0.93	0.74	.48	1.12	-	-	-	-	-
B212	0.33	0.81	0.67	.52	1.15	-0.95	0.93	0.84	.43	1.40
B215	0.27	0.90	0.67	.53	1.19	-	-	-	-	-
A051	0.17	1.01	0.84	.49	1.11	-1.11	0.95	0.91	.46	0.74
B153	0.15	0.99	0.97	.45	1.02	-	-	-	-	-
A107	0.14	0.93	0.73	.51	1.11	-0.95	0.97	0.86	.43	1.20
A098	0.10	1.04	1.04	.43	0.97	-	-	-	-	-
A105	0.09	0.96	0.88	.48	1.05	-2.65	0.80	0.69	.23	0.89
B174	0.09	1.36	1.21	.43	1.02	-	-	-	-	-
A067	0.03	1.03	0.94	.49	1.07	-0.63	0.93	0.86	.52	1.19
A068	-0.01	1.40	1.50	.38	0.86	-	-	-	-	-
B158	-0.03	0.96	0.84	.5	1.07	-0.90	0.95	0.95	.61	0.67
B175	-0.06	0.77	0.69	.56	1.14	-	-	-	-	-
B122	-0.12	1.09	0.97	.48	1.02	-	-	-	-	-
B176	-0.14	0.96	0.91	.51	1.06	-0.83	0.96	0.95	.57	0.63
A014	-0.22	1.15	1.19	.42	0.85	-0.60	1.00	1.00	.6	0.20
A104	-0.30	1.01	0.98	.47	0.91	-	-	-	-	-
B213	-0.36	0.98	1.03	.47	0.88	-	-	-	-	-
A052	-0.37	1.27	1.25	.44	0.84	-0.71	0.98	0.99	.56	0.29
B155	-0.39	1.01	1.17	.43	0.76	-	-	-	-	-
A106	-0.50	0.99	1.01	.5	0.91	-	-	-	-	-
A053	-0.65	1.04	1.13	.48	0.76	-0.52	0.97	0.97	.42	0.66
B172	-0.74	0.96	0.99	.56	0.99	-	-	-	-	-
A005	-0.76	1.14	1.5	.36	0.28	-	-	-	-	-

Note: The lines corresponding to excluded items from the GGUM are represented with a hyphen.

When comparing the indices of difficulty for the 12 common items analyzed in the two models we verified that only one item (A053) presented a higher level of difficulty in the unfolding model, which shows that in this model the items are mostly assessed as easy. This fact may be the result of the logic that underlies the GGUM model, which, in contrast to the model of graduate response, is based strictly on the subjects and the latent construct. For the GGUM model this assumption seems to be equally functional or even sufficient for the estimation of the model parameters (Roberts et al., 2000).

When the theta-item correlations obtained are compared, it appears that for seven items the magnitudes of the correlations were higher in the graduated response model, and for five items these magnitudes were higher in the GGUM. Item B123 presented a correlation close to zero (negative) in the unfolding model, which is evidence that the item does not help to assess the theta. Although no items showed equal magnitude, two items (A051 and A067) had a difference lower than .05 (which can be seen by subtracting the theta-total correlations between the two models).

In the model of graduated answer, the discrimination is continuous for all items, being attributed value 1; empirically, however, discriminations do not present this value. The Winsteps software may provide, by means of a post-hoc analysis, the estimative of the discrimination of each item (Linacre, 2009). Analyzing the parameter of discrimination of the items, still taking into consideration the 12 items analyzed in both models, we observed that most items (N=8) presented greater discriminative power in the analysis by graduated response. An exception to that was item B123, in the GGUM, with a quite elevated, power of discrimination; however, the outfit of the item was lower than 0.05, and the correlation with theta was close to zero, which may evidence a distortion of the model.

The indices analyzed in this study provide information indicating a better adequacy of the aggression dimension items to the graded response model. This consideration can be made when comparing the 12 items that were analyzed in the two models. This statement is even more relevant when one considers that, based on the GGUM, the full set of items could not be analyzed to meet the assumptions of the model (i.e., similar levels of standard error), thus reducing the amount of information obtained by the model when compared to the unfolding model (by Winsteps in this case), whereby the set of items as a whole was analyzed.

Thus, these results corroborate what is presented in the literature (Wang et al., 2013), that the unfolding model should work more adequately for items whose content tends to neutrality and not to an extreme, there was a more impaired suitability of this model compared to the most traditionally used dominance model. Together with other evidence (e.g., the level of difficulty of the items compared to the average theta of participants), the unsuitability of sets of items for the unfolding model may be an indicative of validity for tests that explicitly seek to assess an extreme of a specific construct. Even so, we did not find any study in the literature using the application of this model for such purpose.

Final Considerations

In this study, we investigated the suitability of the dominance model (represented by the model of graduated response) and the unfolding model (represented by the GGUM) for the Aggressiveness dimension of the IDCP. The results demonstrated that the dominance model presented a better fit. The findings by Stark et al. (2006), Chernyshenko et al. (2001), and Meijer and Baneke (2004) indicated that the items of a single peak are often neutral and are located in the middle of the continuum of the latent trait.

In these items, the unfolding model may add a better accuracy of measurement in an area where the dominance items are hard to formulate.

The items assessed in the present study correspond to a dimension that evaluates pathological aspects of personality, which places them at an extreme area of the construct; that is to say, part of the items do not tend towards neutrality (i.e., not assessing extremes traits), which may explain the better fit of the items to the dominance model. Therefore, although recent studies suggest some advantages in the use of unfolding models (Chernyshenko et al., 2001; Chernyshenko et al., 2007; Meijer & Baneke, 2004; Stark et al., 2006), the results found in the present study seem to indicate that the unfolding model does not offer advantages in the assessment of items that measure pathological aspects of personality. However, we should note that after the exclusion of the items based on standard errors, only 12 items were used in the analyses, which may bias the estimation of the parameters, and which may also explain a greater inadequacy of the indices found with the unfolding model.

We also want to highlight that further studies with other instruments, or even studies with other dimensions of the IDCP, are necessary to obtain empirical evidence of the results of the present study. Further studies should investigate the reproducibility, with other sets of items, of the pattern found. Furthermore, it would be interesting to investigate the flexibility of the unfolding models in the face of items with fewer trends towards neutrality.

References

APA - American Psychiatric Association. (2002). *Diagnostic and statistical manual of mental disorders – DSM-IV-TR* (4th ed.). Washington, DC: American Psychiatric Association.

Baker, F. B. (2001). *The basics of item response theory*. Washington: ERIC Clearinghouse on Assessment and Evaluation.

Bortolotti, S. L. V. (2010). *Resistência à mudança organizacional: medida de avaliação por meio da teoria da resposta ao item* (Tese de doutorado não publicada). Universidade Federal de Santa Catarina, Florianópolis, Santa Catarina, Brasil.

Carvalho, L. de F. (2011). *Desenvolvimento e verificação das propriedades psicométricas do Inventário Dimensional Clínico da Personalidade* (Tese de doutorado não publicada). Universidade São Francisco, Itatiba, Brasil.

Carvalho, L. de F. & Primi, R. (2011). *Inventário Dimensional Clínico da Personalidade*. Teste psicológico não publicado.

Chernyshenko, O. S., Stark, S., Chan, K. Y., Drasgow, F., & Williams, B. (2001). Fitting item response theory models to two personality inventories: Issues and insights. *Multivariate Behavioral Research, 36*(4), 523-562. doi: 10.1207/S15327906MBR3604_03

Chernyshenko, O. S., Stark, S., Drasgow, F., & Roberts, B. W. (2007). Constructing personality scales under the assumptions of an ideal point response process: Toward increasing the flexibility of personality measures. *Psychological Assessment*, *19*(1), 88-106. doi: 10.1037/1040-3590.19.1.88

Conn, S. R. & Rieke, M. L. (1994). *16PF fifth edition technical manual*. Champagne, IL: Institute for Personality and Ability Testing, Inc.

Coombs, C. H. (1964). *A theory of data* (1st ed.). New York: John Wiley & Sons, Inc.

Embretson, S. E. (2000). Multidimensional measurement from dynamic tests: Abstract reasoning under stress. *Multivariate Behavioral Research, 35*(4), 505-542.

Embretson, S. E. & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: L. Erlbaum Associates.

Linacre, J. M. (2009). Winsteps (Version 3.68) [Computer Software]. Beaverton, Oregon: Winsteps.com.

Linacre, J. M. & Wright, B. D. (1994). Chi-square fit statistics. *Rasch Measurement Transaction, 8*(2), 261-360.

Meijer, R. R. & Baneke, J. J. (2004). Analyzing psychopathology items: A case for nonparametric item

response theory modeling. *Psychological Methods, 9*(3), 354-368. doi: 10.1037/1082-989X.9.3.354

Roberts, J. S., Donoghue, J. R., & Laughlin, J. E. (1999, april). *Estimability of parameters in the generalized graded unfolding model*. Presented at the annual meeting of the American Educational Research Association, Montreal, Canada.

Roberts, J. S., Donoghue, J. R., & Laughlin, J. E. (2000). A general item response theory model for unfolding unidimensional polytomous responses. *Applied Psychological Measurement*, *24*(1), 3-32. doi: 10.1177/01466216000241001

Stark, S., Chernyshenko, O. S., Drasgow, F., & Williams, B. A. (2006). Examining assumptions about item responding in personality assessment: Should ideal point methods be considered for scale development and scoring? *Journal of Applied Psychology*, *91*(1), 25-39. doi: 10.1037/0021-9010.91.1.25

Wang, W., Tay, L., & Drasgow, F. (2013). Detecting differential item functioning of polytomous items for an ideal point response process. *Applied Psychological Measurement*, *37*(4), 316-335.

Wright, B. D. & Stone, M. H. (2004). *Making measures*. Chicago: Phaneron Press, Incorporated.