

Psychopathy Checklist: Youth Version psychometric properties in an Item Response Theory polytomous model

Propriedades psicométricas da Psychopathy Checklist: Youth Version em um modelo politômico da Teoria da Resposta ao Item

Raul Corrêa **FERRAZ**¹  0000-0002-6292-6111

Fernando de Jesus **MOREIRA JUNIOR**²  0000-0001-7863-8424

Fernanda de **VARGAS**³  0000-0003-4984-3503

Fernanda Xavier **HOFFMEISTER**⁴  0000-0003-1932-3259

Gabriel José Chittó **GAUER**³  0000-0003-3092-0185

Silvio José Lemos **VASCONCELLOS**⁴  0000-0001-6415-7494

Abstract

This study assessed the applicability of the Psychopathy Checklist: Youth Version in a sample of teenagers confined in socio-educational institutions. Using an Item Response Theory approach, item properties of this instrument were reviewed using the generalized partial credit model. Eight of the original twenty items of the original instrument were discarded due to low discrimination parameters. As expected, the most discriminating items in the assessment of psychiatric traits

▼ ▼ ▼ ▼ ▼

¹ University of South Carolina, College of Arts and Sciences, Department of Psychology, Pendleton Street, 1512, Barnwell College, 29208, Columbia, SC, United States. Correspondence to: R.C. FERRAZ. E-mail: <raulc Ferraz@gmail.com>; <rferraz@email.sc.edu>.

² Universidade Federal de Santa Maria, Centro de Ciências Naturais e Exatas, Departamento de Estatística. Santa Maria, RS, Brasil.

³ Pontifícia Universidade Católica do Rio Grande do Sul, Escola de Direito, Programa de Pós-Graduação em Ciências Criminais. Porto Alegre, RS, Brasil.

⁴ Universidade Federal de Santa Maria, Centro de Ciências Sociais e Humanas, Departamento de Psicologia. Santa Maria, RS, Brasil.

▼ ▼ ▼ ▼ ▼

How to cite this article

Ferraz, R. C., Moreira Junior, F. J., Vargas, F., Hoffmeister, F. X., Gauer, G. J. C., & Vasconcellos, S. J. L. (2020). Psychopathy Checklist: Youth Version psychometric properties in an Item Response Theory polytomous model. *Estudos de Psicologia* (Campinas), 37, e180076. <http://dx.doi.org/10.1590/1982-0275202037e180076>



were those which affective characteristics are more typical in the description of psychopathic traits, and their larger variability among juveniles is reflected in the checklist's answers. Item anchoring, in turn, determined five anchor levels. Conclusions based on the results are twofold: (a) a shorter version of this measure can offer the same level of information obtained from the full instrument and (b) the measure provides more information on average latent trait levels and is inadequate for clinical use.

Keywords: Psychological assessment; Psychometrics; Item Response Theory; Antisocial personality disorder.

Resumo

Este estudo investigou a aplicabilidade da Psychopathy Checklist: Youth Version em uma amostra de adolescentes em cumprimento de medidas socioeducativas. Com base na Teoria de Resposta ao Item, as propriedades dos itens desse instrumento foram analisadas por meio do Modelo de Crédito Parcial Generalizado. Oito dos vinte itens do instrumento original foram eliminados devido a baixos parâmetros de discriminação. Conforme esperado, os itens mais discriminantes na avaliação psiquiátrica foram aqueles de características afetivas típicas de traços de psicopatia, assim como sua maior variação reflete-se nas respostas a tais itens. A ancoragem de itens, por sua vez, permitiu a definição de cinco níveis âncora. Os resultados sugerem duas conclusões: (a) uma versão mais curta do instrumento pode oferecer níveis de informação similares aos do teste completo e (b) o instrumento é mais informativo em níveis médios de traço latente, sendo inadequado para uso clínico.

Palavras-chave: Avaliação psicológica; Psicometria; Teoria de Resposta ao Item; Transtorno da personalidade antissocial.

Personality disorders are a persistent pattern of behavior that deviates significantly from an individual's culture expectations, is diffuse and rigid (American Psychiatric Association, 2013). Psychopathy is considered a serious personality disorder characterized by behavioral and affective problems that can generate a greater tendency towards criminality, although such acts do not, by themselves, define the disorder (Zepinic, 2017). In general, such behaviors tend to be directly associated with a difficulty in generating adequate empathic reactions in an interpersonal context (Robinson & Rogers, 2015). Current findings also suggest that violence committed by psychopaths tends to be more instrumental and therefore geared to personal advantages rather than merely reactive (Dhingra & Boduszek, 2013).

Although this disorder can only be diagnosed in adulthood, adolescents with psychopathic traits show a heightened sense of grandiosity, self-centeredness, manipulative ability, emotional insensitivity, lack of empathy, lack of remorse and, irresponsibility among other antisocial tendencies (Vahl, Colins, & Lodewijks, 2014). High levels of antisocial traits in adolescents are also associated with violent acts and episodes in which the individual is often in conflict with the law (Loeber, Burke, & Pardini, 2009; Salekin & Frick, 2005).

One of the tools developed for the diagnosis of psychopathy is the Psychopathy Checklist-Revised (PCL-R), developed by Robert Hare and considered the "gold standard" in the evaluation of psychopathy (Vitacco, Neumann, & Jackson, 2005). This instrument aims to comprehensively evaluate the different symptoms that characterize psychopathy, measuring affective, interpersonal and behavioral characteristics. In general terms, this checklist brings together twenty separate items, scored by means of a three-point ordinal scale considering, to this end, the degree to which the subjects' behavior and personality match the description presented in the manual (Hare, 2003). In Brazil, the PCL-R was adapted from Morana's study (2004).

However, controversies about the most appropriate factorial structure for the PCL-R have been recurrent after its use in different cultures and in different countries (Vitacco, Neumann, Caldwell, Leistico, & Van Rybroek, 2006). A systematic review performed by Hauck Filho, Teixeira, and Almeida (2014), evidenced, for example, that the two-factor model is still recurrent in recent scientific literature, although it is not the only one mentioned in different studies that used the confirmatory factor analysis for this same check list.

The observation of the items in the Psychopathy Checklist: Youth Version (PCL: YV) reveals that many items came directly from the PCL-R, with little or no change. International studies evaluate the validity and reliability of the instrument as appropriate (Forth, Kosson, & Hare, 2003), but its predictive ability in relation to criminal recidivism has been scrutinized (Olver, Stockdale, & Wormith, 2009; Shepherd & Strand, 2016), and comorbidities may also decrease the predictive value of instruments, including the PCL: YV (Khanna, Shaw, Dolan, & Lennox, 2014). In men, the second factor (“socially deviant lifestyle”) is predominantly responsible for the predictive value. In women an equivalence is inferred (Cope, Ermer, Nyalakanti, Calhoun, & Kiehl, 2014), despite the lack of truly conclusive data in this regard (Walters, 2014).

In the McCuish, Mathesius, Lussier, and Corrado (2018) study with Canadian natives and young Caucasians, PCL: YV outcomes were similar between the two groups, although the first group was overrepresented in the penal system. However, in the work of Sitney, Caldwell, and Caldwell (2016), African Americans revealed significantly higher scores and recidivism rates when compared to young Caucasians, even when controlling for other influences. In particular, in cultures other than the North American, there is still a significant lack of such studies (Pechorro, Barroso, Maroco, Vieira, & Gonçalves, 2015). Regarding the association between PCL: YV scores with other variables, the data obtained with this instrument are similar to those suggested from studies with the adult version, including the relationship between symptoms of psychopathy with drug abuse (Vincent, Cope, King, Nyalakanti, & Kiehl, 2018).

The development of more reliable measures to assess psychopathic traits in young people has shown that these same traits tend to be more present in individuals with chronic criminal trajectories (McCuish, Corrado, Lussier, & Hart, 2014; Tsang, 2018), as well as other risk factors (Chauhan et al., 2014). In addition, cognitive mediating aspects between scoring and future violent behavior also need to be better understood (Lennox & Dolan, 2014; Kahn, Ermer, Salovey, & Kiehl, 2016; Walters & DeLisi, 2015).

Method

Participants

The investigation sample consisted of 255 male adolescents in conflict with the law, evaluated while complying with socio-educational measure of deprivation of liberty or provisional detention. Data were collected during the period from 2008 to 2010.

The inclusion criteria used were: age between 12 and 19 years old; being in socio-educational condition of deprivation of liberty or in provisional detention at *Fundação de Atendimento Socioeducativo* (FASE, Socio-Educational Care Foundation); voluntarily agreeing to participate in the investigation. On the other hand, exclusion criteria were: active psychotic signs or symptoms, intellectual disability or hearing impairment clinically detected or described in the medical records; refusal to complete the application of the instruments at any stage of the process.

Instruments

Hare’s Psychopathy Inventory: Youth Version (Gauer, Vasconcellos, & Werlang, 2006) includes an interview guide, plus the technical manual and a response form to be completed by the evaluator, consisting of 20 items that investigate affection, interpersonal aspects and lifestyle. Brazilian studies, although incipient, indicate adequate inter-rater reliability for PCL: YV (Ronchetti, Davoglio, Salvador-Silva, Vasconcellos, & Gauer, 2010). The instrument is scored considering the adolescent’s usual functioning since late childhood/early adolescence except for one item that also considers the functioning prior to the first decade of life. Each item

is scored on a three-point grading: zero (0) characteristics are not present, or oppose, or are inconsistent with the intent of the item; one (1) when the item partially applies but not to the degree required for a higher score; or when there is uncertainty about the application of the item, or there are conflicts between sources of information that cannot be resolved in favor of 0 or 2; two (2) when the item applies perfectly in its most fundamental aspects. Exceptionally, when there is not enough information to score an item, the item's score may be omitted, but omissions may never exceed five out of twenty.

Data Analysis

The objectives of the analyses performed are: to obtain the PCL: YV item parameters in a generalized partial credit model, and to build a scale using anchor levels established from the selected items.

Item Response Theory

Although it includes behavioral aspects, the construct called psychopathy cannot, of course, be directly observed. In this connection, statistical models of latent traits (variables that cannot be directly observed but inferred from observable variables) are one of the main objects of investigation for quantitative methods in social sciences. In particular, the Item Response Theory (IRT) proposes to connect a person's probability of correctly answering an item with a combination of that person's ability and the attributes of that item (Andrade, Tavares, & Valle, 2000). For Pasquali (2003), IRT adopts two fundamental axioms: 1) Performance on a test item is the effect and latent traits are the cause; 2) The performance-latent trait relationship can be described by an increasing monotonic equation called the Item Characteristic Curve (ICC).

Interpretation

The interpretation of the parameters is made as follows: for Ayala (2009), a "reasonably good" discrimination parameter (α) varies approximately between 0.8 and 2.5. Baker and Kim (2017) claim as moderate α between 0.65 and 1.34. The α discrimination of each item reaches its maximum value in the difficulty level of this item's ability. The difficulty parameter β corresponds to the point on the ability scale where $P(\theta) = 0.5$; when an item is easy; this occurs at a low skill level; when it is difficult, it occurs at a high level. These skill levels are usually set out as a function of the standard deviation from the mean on the scale (0, 1), and then each level is given a qualitative interpretation (so that data presentation becomes more pedagogical) and are then called *anchor levels*.

The number of anchor levels is not known *a priori*. To determine at which level an item is anchor (or even if an item is anchor at some level), Beaton and Allen (1992) indicate the following conditions: 1) positive answer by a large percentage of individuals (at least 65%) with this skill level; 2) be answered positively by a smaller percentage of individuals (maximum 50%) with the immediately previous skill level; 3) the difference between these two proportions must be at least 0.30.

Model selection

PCL: YV is an instrument with polytomic items, requiring consistent modeling. Choosing not to consider the hypothesis of items with the same discrimination (the case of Rasch models), one has to choose between two polytomic models: the Generalized Partial Credit Model (GPCM) (Masters, 1982; Muraki, 1992) or the

Graded Response Model (GRM) (Samejima, 1968). According to Tsang, Piquero, and Cauffman (2014), although appropriate for categorical responses, GRM assumes that the likelihood of responding to a higher category increases as the latent trait increases.

In fact, this is expected. However, this hypothesized ordering may not be reflected in the empirical data. In polytomous items, if the probabilities do not increase along with the latent trait, at least one response category will never be the most likely (for any point on the θ scale; Andrich, 1988). The assumption of ordered categories made by the GRM implies that this model is unable to detect these “disorganized thresholds” (Tsang et al., 2014). The GPCM, while making the same assumption, does not force the thresholds to be ordered.

In the interpretation of the discrimination parameter the criterion of Baker and Kim (2017), accepting as moderate or higher an $\alpha \geq 0.65$ was adopted. The α discrimination of each item reaches its maximum value at the difficulty level of the item. The difficulty parameter β of the item corresponds to the point on the ability scale where $P(\theta) = 0.5$: when an item is easy, this occurs at a low skill level; when it is difficult, it occurs at a high level. In polytomous models, as in the case of this paper, each score of an item (as in the case of PCL: YV; 0, 1, and 2) is called a category and exhibits its own characteristic curve. The intersection parameter δ_j of one category with the previous one indicates the latent trait levels where it is equally likely to answer either score (0 or 1, 1 or 2). This parameter refers to the exact points of intersection, not to the category as a whole. An item with m_j categories exhibited m_{j-1} intersection points. As Muraki (1992) arbitrarily defined $\delta_1 = 0$, we work with δ_2 as an intersection between categories 0 and 1, and δ_3 as an intersection between 1 and 2, in the case of PCL: YV. In the interpretation of anchor levels, since there were few items left, the criteria 2 and 3 previously described were relaxed: percentages of up to 55% in the previous skill level were accepted, as well as differences greater than 0.25, a procedure also adopted by Mafra (2011).

Results

Parameters

The study with polytomous data considered the twelve items that presented discrimination parameter (α) higher than 0.65. The categories δ of the items, reviewed by the generalized partial credit model, referred to a three-point Likert Scale, covering the scores “0”, “1” and “2” equivalent to “The item does not apply to the individual”, “The item applies to some degree, but not to the point of representing a score 2, or “there are many conflicting exceptions or information” and “The item reasonably applies to the individual in the most essential aspects”, respectively. Table 1 presents (a) the discrimination and location parameters of the categories of all items, considering the application in GPCM by the R Software (R Core Team, 2013), through the *ltm* package (Rizopoulos, 2006); and (b) the recalculated parameters of the remaining items after the elimination of the items with low discrimination (1, 2, 4, 13, 14, 16, 17 and 19).

Figure 1 presents the characteristic curves of the categories of the 12 selected items. Items 8, 6, and 18 presented the greatest discrimination parameters (α s = 3.00, 2.00, and 1.75, respectively), resulting in steeper curves and suggesting that they are more sensitive to different levels of psychopathy traits than items 9, 3 and 11, which presented the lowest α s among the selected items, for example (0.71, 0.74 and 0.80, respectively).

The data in Table 1 and Figure 1 also show the difficulty parameters of the categories and the latent trait range that each test response is most likely to meet. For example, in item 9, “Parasitic Lifestyle”, a score 2 answer (“The item applies reasonably well to the individual in the most essential aspects”) only becomes the most likely answer from point 1.89 on the latent trait scale (and yet it is only close to 0.5); on the other hand, an individual with the same θ approaches the probability of 1 for a response of item 6, “Lack of Remorse/Guilt”.

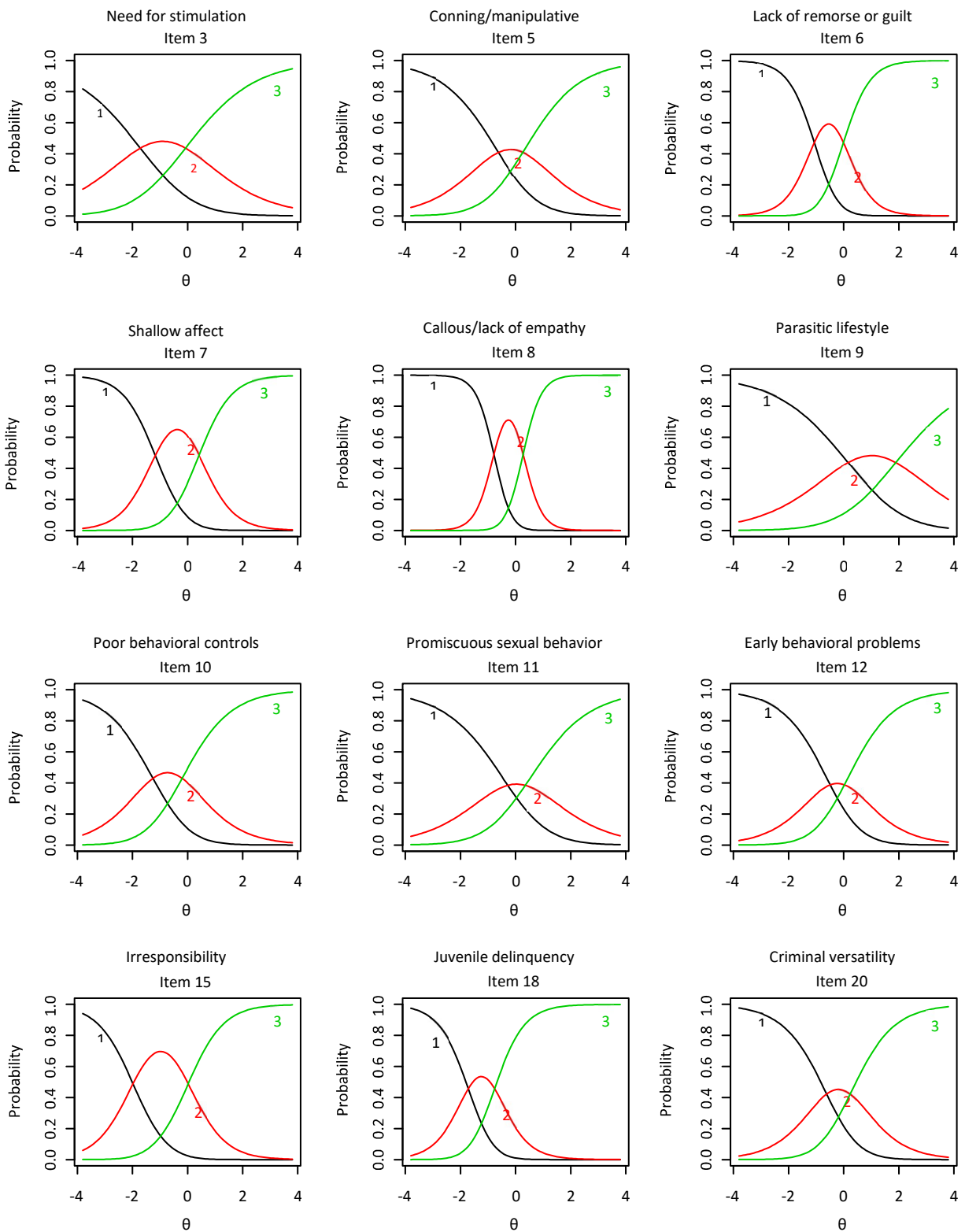


Figure 1. Psychopathy Checklist: Youth Version Item Response Curves in the reestimated model.
 Note: θ (theta): latent trait.

Table 1

Estimated item parameters in two models in the Generalized Partial Credit Model

Items	α	α_R	δ_2	δ_{2R}	δ_3	δ_{3R}
1	0.364		-3.077		0.020	
2	0.496		-1.991		0.134	
3	0.764	0.748	-1.690	-1.712	-0.064	-0.075
4	0.446		-1.768		0.769	
5	0.919	0.902	-0.618	-0.625	0.267	0.259
6	1.758	2.005	-1.087	-1.068	-0.020	-0.006
7	1.613	1.655	-1.176	-1.171	0.425	0.415
8	2.703	3.002	-0.794	-0.791	0.280	0.275
9	0.801	0.716	0.132	0.161	1.777	1.892
10	1.019	1.054	-1.264	-1.255	-0.204	-0.199
11	0.809	0.802	-0.277	-0.283	0.359	0.353
12	1.046	1.063	-0.478	-0.486	0.029	0.028
13	0.604		-1.272		0.789	
14	0.605		-2.231		-1.023	
15	1.642	1.522	-1.920	-1.977	0.039	0.025
16	0.460		-0.291		0.666	
17	0.605		-1.011		1.310	
18	1.828	1.753	-1.703	-1.715	-0.738	-0.764
19	0.552		2.204		-0.244	
20	1.252	1.176	-0.610	-0.616	0.257	0.237

Log.Lik: original = -4663.968, reestimated = -2690.084

Note: α (alpha): discrimination parameter. δ (delta): intersection parameter. Log.Lik: model fit statistic: likelihood logarithm to convergence of the model parameter estimation algorithm. R: parameters reestimated after exclusion of the items with low discrimination, composing a second model. Bold values: $\alpha < 0.65$.

Figure 2 presents the Item Information Functions (IIFs); the IIF peaks indicate where an individual with a latent trait truly at that level can be accurately estimated (Baker & Kim, 2017). The smaller the amount of information, the lower the accuracy of θ estimates. Item 3, for example, is more accurate (with information close to 15) in the latent trait range between -1 and 0, and loses accuracy as it approaches extremes. Figure 2 also shows at which points on the latent trace scale the test provides more accurate estimates on individuals. It is clear that the test is positioned to the left in relation to the mean and that most of the information is located approximately between -1 to 1 on the latent trace scale. The test provides little information about respondents located outside this range, which are in significant numbers and whose interviews represent a considerable spending of time in prison surveys. In summary, these results suggest two conclusions: (a) the short version of the instrument offers similar levels of information as the full test and (b) the instrument is more informative at medium latent trace levels and is not indicated for use in the clinical setting.

Anchors

As stated above, by definition $\delta_1 = 0$: the probability of obtaining a score "0" or higher is always 1, regardless of the position on the scale. After calculating the cumulative probabilities for δ_{2R} and δ_{3R} (the probability of obtaining the score "1" or "2" at each level of the latent trace scale) and only of the upper category δ_{3R} (score "2"), it was possible to establish five anchor levels: -1, 0, 1, 2 and 3.

Table 2 shows a synthesis of the anchor items for the categories equivalent to the scores "1" and "2", according to the values found for δ_{2R} and δ_{3R} , and the relevant levels that determine anchor categories.

According to Table 2, individuals at level -1 are characterized for achieving a score of "1" or higher for the items "Need for Stimulation", "Parasitic lifestyle" and "Irresponsibility", which fall into the Behavioral

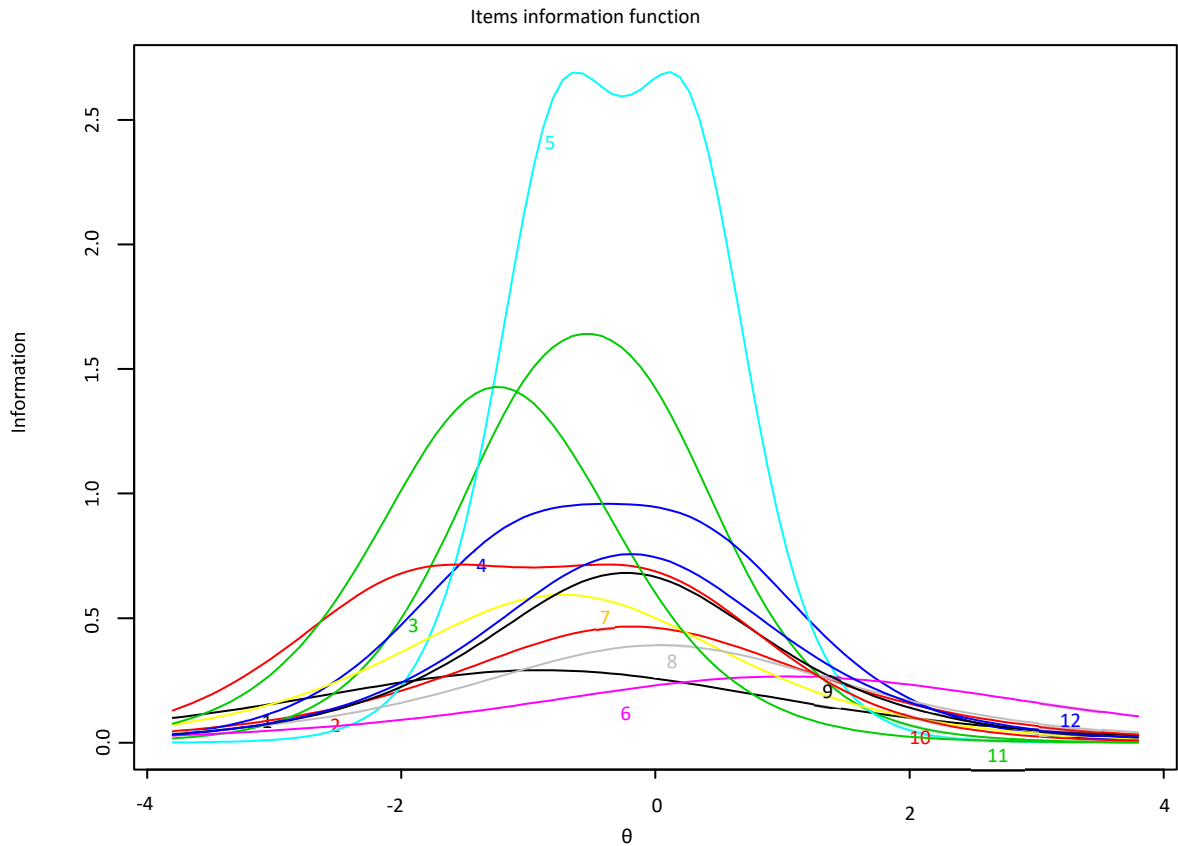


Figure 2. Item Information Function in the reestimated model.
 Note: θ (theta): latent trait.

Table 2

Classification of anchor categories of items and assessment factor by latent trait level

1 of 2

Item	Description	Group
Level -1		
Item 3 (δ_{2R})	Need for stimulation – ‘1’	Behavioral
Item 9 (δ_{2R})	Parasitic lifestyle – ‘1’	Behavioral
Item 10 (δ_{2R})	Poor behavioral controls – ‘1’	Antisocial
Item 15 (δ_{2R})	Irresponsibility – ‘1’	Behavioral
Item 18 (δ_{2R})	Juvenile delinquency – ‘1’	Antisocial
Level 0		
Item 5 (δ_{2R})	Conning/manipulative – ‘1’	Interpersonal
Item 6 (δ_{2R})	Lack of remorse or guilt – ‘1’	Affect
Item 7 (δ_{2R})	Shallow affect – ‘1’	Affect
Item 8 (δ_{2R})	Callous/lack of empathy – ‘1’	Affect
Item 11 (δ_{2R})	Promisc. Sex Behavior – ‘1’	Interpersonal
Item 12 (δ_{2R})	Early behav. problems – ‘1’	Antisocial
Item 18 (δ_{3R})	Juvenile delinquency – ‘2’	Antisocial
Item 20 (δ_{2R})	Criminal versatility – ‘1’	Antisocial
Level 1		
Item 3 (δ_{3R})	Need for stimulation – ‘2’	Behavioral
Item 6 (δ_{3R})	Lack of remorse or guilt – ‘2’	Affect
Item 7 (δ_{3R})	Shallow affect – ‘2’	Affect
Item 8 (δ_{3R})	Callous/lack of empathy – ‘2’	Affect

Note: δR (delta): reestimated intersection parameter.

Table 2

Classification of anchor categories of items and assessment factor by latent trait level

2 of 2

Item	Description	Group
Item 10 (δ_{3R})	Poor behavioral controls – '2'	Antisocial
Item 12 (δ_{3R})	Early behav.problems – '2'	Antisocial
Item 15 (δ_{3R})	Irresponsibility – '2'	Behavioral
Item 20 (δ_{3R})	Criminal versatility – '2'	Antisocial
Level 2		
Item 11 (δ_{3R})	Promisc. Sex Behavior – '2'	Interpersonal
Level 3		
Item 9 (δ_{3R})	Parasitic lifestyle – '2'	Behavioral

Note: δR (delta): reestimated intersection parameter.

group, and also in the items "Poor behavioral controls" and "Severe criminal behavior", which fall into the Antisocial group. Likewise, in the following levels, the categories of each item that correspond to certain items are successively exposed. It is important to note that this does not mean that a respondent of any level will necessarily respond in this exposed pattern, since these are probabilities: anchor levels have the function of allowing a more accessible and pedagogical interpretation of the data.

Table 2 also shows that the vast majority of item categories are characterized as providing more information about individuals between levels -1 and 1. Only the δ_{3R} categories of the items "Promiscuous sexual behavior" and "Parasitic lifestyle" define levels 2 and 3 respectively. It is an issue that a complex-use instrument such as PCL: YV, which requires a lot of time to apply and to train the surveyors, has such a limited range of information. In fact, this may represent an under-representation of precisely the portion of the population that requires more attention from mental health professionals in socio-educational institutions.

Discussion

The properties and functioning of the PCL: YV items in a sample of male adolescents in the state of *Rio Grande do Sul* were presented here. The results of the reviews performed through the IRT suggest that out of the 20 items of the original version of the instrument, 12 presented satisfactory discrimination parameters. It is noteworthy that the lack of adjustment of the other 8 items may be due to the few answers in those categories, considering the limited sample. Some items are relatively better at assessing different levels of psychopathy (for example, items 6 and 8: "lack of remorse/guilt" and "callous/lack of empathy"); in fact, these affective characteristics can be considered traditional in the description of psychopathy traits, and their greater variability among young people is reflected in the responses to these items. Other items, such as "Need for Stimulation", were less sensitive: once again, this may be due to a greater homogeneity of this trait among adolescents.

Some items were more extreme, i.e., a higher level of latent trait is required to be scored with a "The item applies to the individual reasonably well in the most essential aspects" response by the surveyor. Two items that exhibited this functioning includes "Parasitic Lifestyle" and "Promiscuous Sexual Behavior," which is consistent with the manual information and PCL: YV training. Considering that most adolescents are financially dependent on their family, as well as the social vulnerability to which the target audience of the socio-educational institutions is exposed to, the task of the evaluator is made more difficult and a more extreme trait becomes necessary to be given a high score. Similarly, the item "Promiscuous sexual behavior"

also presents particular challenges in the assessment: a teenager who has never engaged in sexual intercourse receives a score of zero, and even if he or she has been involved, it is likely that only the individuals with the highest latent trait levels will be willing to talk about these experiences and describe them as impersonal or trivial.

It can be seen from the results that the instrument offers more information in the near-average latent trait range. The same causes that affect the parameters of discrimination and difficulty discussed earlier also affect the performance of the items that were maintained. However, response distortions that may affect the behavior of items should also be considered. Response distortions are systematic tendencies to respond to a range of items that are not relevant to the construct being measured (Bandalos, 2018), resulting in biases that affect instrument accuracy and validity. Social desirability is the most well-known and studied distortion and comes in different forms: two of its components that we hypothesize to have a significant influence in this study are self-deception enhancement, in which the respondent perceives himself more favorably than he ought to, and impression management, when a conscious effort to conceal “flaws” exists. Other relevant distortion in teenager evaluation is acquiescence, the tendency to disagree or agree with most items regardless of their content. Response distortion is not a trivial problem in psychological assessment and is particularly challenging in antisocial behaviors investigations: in future studies, we suggest the use of instruments that can detect patterns of social desirability, including other scales and implicit association tests.

In the psychometric context, the construction and comparison of different statistical models is an essential step in the development of tests and in obtaining evidence of construct validity - or, in a more modern terminology, internal structure validity (Messick, 1995; American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014; Bandalos, 2018). In the case of IRT applications, comparison between studies is possible provided that the formulation of the models involved is respected. The following explanation is brief and therefore superficial, but technical details are available in the psychometric literature (e.g., Maydeu-Olivares, 2005). Using the Thissen and Steinberg (1986) taxonomy, there are two major classes of parametric models in IRT: difference models and division by total models. The main difference model is Samejima’s GRM (1968). Among the division by total models, the most used in applications are those that do not include a guessing parameter; three of them are considered nested (i.e., the simpler models can be obtained by limiting the more general model): Bock’s nominal model (1972), Thissen and Steinberg’s ordinal model (1986) and Masters’ partial credit model (1982). The generalized partial credit model is equivalent to the formulation of Thissen and Steinberg. In the case of the present study, the estimated parameters can be compared with the outcome of other studies that used nested models; this same comparison cannot be done with GRM, for example (in fact, GRM and GPCM have the same number of parameters). However, it is noteworthy that GRM and GPCM are not equivalent, allowing the comparison of model fit statistics.

Leaving aside cultural differences, comparison with English-language work that also used the GPCM can also be helpful. Some of the items in this study with higher discrimination parameters also presented good parameters in the study by Tsang et al. (2014), such as “Callous/Lack of Empathy” and “Lack of Remorse/Guilt”, suggesting that these attributes are good indicators for differentiating between different levels of psychopathy traits. Similarly, items such as “Need for Stimulation” also proved to be poorly discriminating in other studies (Tsang et al., 2015, 2014), suggesting that this is a more universal feature of adolescence and therefore less informative for the purpose of the PCL: YV. Other items also presented more inconsistent discriminations in these multiple studies, such as, for example, the item “Shallow emotions”. Besides cultural differences that could influence the evaluators’ criteria on what constitutes the demonstration and the experience of emotions, it would be interesting to explore other explanations in studies in Brazil with samples of individuals from different institutions and using Differential Item Functioning.

Final Considerations

Despite the limitations, it is noteworthy that as far as the authors are aware this is the first study in Brazil to evaluate the properties of PCL: YV items. It is important to highlight, however, that these findings are related to this selected sample. Further studies with similar methods with samples from different institutions are encouraged, as well as studies that include Differential Item Functioning analyses for aspects such as age, gender, geographic region, and ethnicity. Caution is also advised in the use of PCL: YV and its scores by practitioners and investigators, who should take into account these potential differences. This study does not support the indiscriminate use of PCL: YV in assessments of young offenders in the southern region or elsewhere in Brazil.

Acknowledgement

The authors are grateful to the anonymous reviewers whose collaborative comments have enhanced this article.

Contributors

R. C. FERRAZ and F. J. MOREIRA JUNIOR designed the study, conducted statistical analyses, wrote and reviewed the manuscript. G. J. C. GAUER and S. J. L. VASCONCELLOS designed the study, coordinated data collection, wrote and reviewed the manuscript. F. VARGAS and F. X. HOFFMEISTER coordinated data collection, wrote and reviewed the manuscript.

References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for Educational and Psychological Testing*. Washington: American Educational Research Assn.
- American Psychiatric Association. (2013). *Manual Diagnóstico e Estatístico de Transtornos Mentais: DSM –V*. (5a ed.). Porto Alegre: Artmed.
- Andrade, D. F., Tavares, H. R., & Valle, R. C. (2000). *Teoria da resposta ao item: conceitos e aplicações*. São Paulo: Associação Brasileira de Estatística. Recuperado de https://docs.ufpr.br/~aanjos/CE095/LivroTRI_DALTON.pdf
- Andrich, D. (1988). A general form of Rasch's extended logistic model for partial credit scoring. *Applied Measurement in Education*, 1(4), 363-378. http://dx.doi.org/10.1207/S15324818AME0104_7
- Ayala, R. J. (2009). *The theory and practice of Item Response Theory*. New York: Guilford Press.
- Baker, F. B., & Kim, S-H. (2017). *The basics of Item Response Theory using R*. New York: Springer. <http://dx.doi.org/10.1007/978-3-319-54205-8>
- Bandalos, D. L. (2018). Methodology in the social sciences. *Measurement theory and applications for the social sciences*. New York: Guilford Press.
- Beaton, A. E., & Allen, N. L. (1992). Interpreting scales through scale anchoring. *Journal of Educational Statistics*, 17(2), 191. <http://dx.doi.org/10.2307/1165169>
- Bock, R. D. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika*, 37(1), 29-51. <http://dx.doi.org/10.1007/BF02291411>
- Chauhan, P., Ragbeer, S. N., Burnette, M. L., Oudekerk, B., Reppucci, N. D., & Moretti, M. M. (2014). Comparing the Youth Psychopathic Traits Inventory (YPI) and the Psychopathy Checklist-Youth Version (PCL-YV) among offending girls. *Assessment*, 21(2), 181-194. <http://dx.doi.org/10.1177/1073191112460271>
- Cope, L. M., Ermer, E., Nyalakanti, P. K., Calhoun, V. D., & Kiehl, K. A. (2014). Paralimbic gray matter reductions in incarcerated adolescent females with psychopathic traits. *Journal of Abnormal Child Psychology*, 42(4), 659-668. <http://dx.doi.org/10.1007/s10802-013-9810-4>

- Dhingra, K., & Boduszek, D. (2013). Psychopathy and criminal behaviour: a psychosocial research perspective. *Journal of criminal psychology, 3*(2), 83-107. <http://dx.doi.org/10.1108/JCP-06-2013-0014>
- Forth, A. E., Kosson, D. S., & Hare, R. D. (2003). *The psychopathy checklist: youth version*. Toronto: Multi-Health Systems.
- Gauer, G. J. C., Vasconcellos, S. J. L., & Werlang, B. G. (2006). *Inventário de psicopatia de hare: versão jovens*. Toronto: Multi-Health Systems.
- Hare, R. D. (2003). *The hare psychopathy checklist revised* (2nd ed.). Toronto: Multi Health Systems.
- Hauck Filho, N., Teixeira, M. A. P., & Almeida, R. M. M. (2014). Estrutura fatorial da escala Psychopathy Checklist - Revised (PCL-R): uma revisão sistemática. *Avaliação Psicológica, 13*(2), 247-256.
- Kahn, R. E., Ermer, E., Salovey, P., & Kiehl, K. A. (2016). Emotional intelligence and callous: unemotional traits in incarcerated adolescents. *Child Psychiatry & Human Development, 47*(6), 903-917. <http://dx.doi.org/10.1007/s10578-015-0621-4>
- Khanna, D., Shaw, J., Dolan, M., & Lennox, C. (2014). Does diagnosis affect the predictive accuracy of risk assessment tools for juvenile offenders: conduct disorder and attention deficit hyperactivity disorder. *Journal of Adolescence, 37*(7), 1171-1179. <http://dx.doi.org/10.1016/j.adolescence.2014.08.008>
- Lennox, C., & Dolan, M. (2014). Temperament and character and psychopathy in male conduct disordered offenders. *Psychiatry Research, 215*(3), 706-710. <http://dx.doi.org/10.1016/j.psychres.2014.01.019>
- Loeber R., Burke J., & Pardini D. A. (2009). Perspectives on oppositional defiant disorder, conduct disorder, and psychopathic features. *Journal of Child Psychology and Psychiatry, 50*(1-2), 133-142. <http://dx.doi.org/10.1111/j.1469-7610.2008.02011.x>
- Mafra, P. M. R. (2011). *Proposta de uma sistemática para a modelagem de risco de crédito sob a perspectiva da Teoria da Resposta ao Item* (Dissertação de doutorado não publicado). Universidade Federal de Santa Catarina. Recuperado de <https://repositorio.ufsc.br/xmlui/handle/123456789/95148>
- Masters, G. N. (1982). A rasch model for partial credit scoring. *Psychometrika, 47*(2), 149-174. <http://dx.doi.org/10.1007/BF02296272>
- Maydeu-Olivares, A. (2005). Further empirical results on parametric versus non-parametric irt modeling of Likert-type personality data. *Multivariate Behavioral Research, 40*(2), 261-279. http://dx.doi.org/10.1207/s15327906mbr4002_5
- McCuish, E. C., Corrado, R., Lussier, P., & Hart, S. D. (2014). Psychopathic traits and offending trajectories from early adolescence to adulthood. *Journal of Criminal Justice, 42*(1), 66-76. <http://dx.doi.org/10.1016/j.jcrimjus.2013.12.004>
- McCuish, E. C., Mathesius, J. R., Lussier, P., & Corrado, R. R. (2018). The cross-cultural generalizability of the psychopathy checklist: youth version for adjudicated indigenous youth. *Psychological Assessment, 30*(2), 192-203. <http://dx.doi.org/10.1037/pas0000468>
- Messick, S. (1995). Validity of psychological assessment: validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*(9), 741-749. <http://dx.doi.org/10.1037/0003-066X.50.9.741>
- Morana, H. C. P. (2004). Escala Hare PCL-R: critérios para pontuação de psicopatia revisados: Versão Brasileira. São Paulo: Casa do Psicólogo.
- Muraki, E. (1992). A generalized partial credit model: application of an EM algorithm. *Applied Psychological Measurement, 16*(2), 159-176. <http://dx.doi.org/10.1177/014662169201600206>
- Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2009). Risk assessment with young offenders: a meta-analysis of three assessment measures. *Criminal Justice and Behavior, 36*(4), 329-353. <http://dx.doi.org/10.1177/0093854809331457>
- Pasquali, L. (2003). *Psicometria: teoria dos testes na Psicologia e na Educação* (5a ed.). Petrópolis: Vozes.
- Pechorro, P., Barroso, R., Maroco, J., Vieira, R. X., & Gonçalves, R. A. (2015). Psychometric properties of the psychopathy checklist: youth version among portuguese juvenile delinquents. *International Journal of Offender Therapy and Comparative Criminology, 59*(12), 1322-1337. <http://dx.doi.org/10.1177/0306624X14535558>
- R Core Team (2013). R: a language and environment for statistical computing. R Foundation for Statistical Computing: Vienna. Retrieved from <http://www.R-project.org/>
- Rizopoulos, D. (2006). ltm: an R package for Latent Variable Modelling and Item Response Theory Analyses. *Journal of Statistical Software, 17*(5), 1-25. Retrieved from <http://www.jstatsoft.org/v17/i05/>
- Robinson, E. V., & Rogers, R. J. (2015). Empathy faking in psychopathic offenders: the vulnerability of empathy measures. *Journal of Psychopathology and Behavioral Assessment, 37*(1), 545-552. <http://dx.doi.org/10.1007/s10862-015-9479-9>
- Ronchetti, R., Davoglio, T. R., Salvador-Silva, R., Vasconcellos, S. J. L., & Gauer, G. J. C. (2010). Inventário de Psicopatia de Hare: Versão Jovens (PCL:YV): estudo preliminar em amostra adolescente brasileira. *Interamerican Journal of Psychology, 44*(3), 540-546.

- Salekin, R. T., & Frick, P. J. (2005). Psychopathy in children and adolescents: the need for a developmental perspective. *Journal of Abnormal Child Psychology*, *33*(4), 403-409. <http://dx.doi.org/10.1007/s10802-005-5722-2>
- Samejima, F. (1968). Estimation of latent ability using a response pattern of graded scores. *ETS Research Bulletin Series*, *1968*(1), i-169. <http://dx.doi.org/10.1002/j.2333-8504.1968.tb00153.x>
- Shepherd, S. M., & Strand, S. (2016). The utility of the Psychopathy Checklist: Youth Version (PCL: YV) and the Youth Psychopathic Trait Inventory (YPI): is it meaningful to measure psychopathy in young offenders? *Psychological Assessment*, *28*(4), 405-415. <http://dx.doi.org/10.1037/pas0000182>
- Sitney, M. H., Caldwell, B. M., & Caldwell, M. F. (2016). The longitudinal relationship between african american status, psychopathic traits, and violent recidivism in juvenile offenders. *Criminal Justice and Behavior*, *43*(9), 1190-1203. <http://dx.doi.org/10.1177/0093854816645619>
- Thissen, D., & Steinberg, L. (1986). A taxonomy of item response models. *Psychometrika*, *51*(4), 567-577. <http://dx.doi.org/10.1007/BF02295596>
- Tsang, S. (2018). Troubled or traumatized youth? the relations between psychopathy, violence exposure, posttraumatic stress disorder, and antisocial behavior among juvenile offenders. *Journal of Aggression, Maltreatment & Trauma*, *27*(2), 164-178. <http://dx.doi.org/10.1080/10926771.2017.1372541>
- Tsang, S., Schmidt, K. M., Vincent, G. M., Salekin, R. T., Moretti, M. M., & Odgers, C. L. (2015). Assessing psychopathy among justice involved adolescents with the PCL: YV: an item response theory examination across gender. *Personality Disorders: Theory, Research, and Treatment*, *6*(1), 22-31. <http://dx.doi.org/10.1037/PER0000094>
- Tsang, S., Piquero, A. R., & Cauffman, E. (2014). An examination of the Psychopathy Checklist: Youth Version (PCL:YV) among male adolescent offenders: an item response theory analysis. *Psychological Assessment*, *26*(4), 1333-1346. <http://dx.doi.org/10.1037/A0037500>
- Vahl, P., Colins, O. F., & Lodewijks, H. P. B. (2014). Psychopathic-like traits in detained adolescents: clinical usefulness of self-report. *European Child & Adolescent Psychiatry*, *23*(1), 691-699. <http://dx.doi.org/10.1007/s00787-013-0497-4>
- Vincent, G. M., Cope, L. M., King, J., Nyalakanti, P., & Kiehl, K. A. (2018). Callous-unemotional traits modulate brain drug craving response in high-risk young offenders. *Journal of Abnormal Child Psychology*, *46*(5), 993-1009. <http://dx.doi.org/10.1007/s10802-017-0364-8>
- Vitacco, M. J., Neumann, C. S., & Jackson, R. L. (2005). Testing a four-factor model of psychopathy and its association with ethnicity, gender, intelligence, and violence. *Journal of Consulting and Clinical Psychology*, *73*(3), 466-76. <http://dx.doi.org/10.1037/0022-006X.73.3.466>
- Vitacco, M. J., Neumann, C. S., Caldwell, M. F., Leistico, A-M, & Van Rybroek, G. J. (2006). Testing factor models of the Psychopathy Checklist: Youth Version and their association with instrumental aggression. *Journal of Personality Assessment*, *87*(1), 74-83. http://dx.doi.org/10.1207/s15327752jpa8701_06
- Walters, G. D. (2014). Predicting self-reported total, aggressive, and income offending with the Youth Version of the Psychopathy Checklist: gender- and factor-level interactions. *Psychological Assessment*, *26*(1), 288-296. <http://dx.doi.org/10.1037/a0034862>
- Walters, G. D., & DeLisi, M. (2015). Psychopathy and violence: does antisocial cognition mediate the relationship between the PCL: YV factor scores and violent offending? *Law and Human Behavior*, *39*(4), 350-359. <http://dx.doi.org/10.1037/lhb0000123>
- Zepinic, V. (2017). Psychopathy: simple or syndromal disorder of personality. *International Journal of Psychological Studies*, *9*(4). <http://dx.doi.org/10.5539/ijps.v9n4p1>

Received: July 20, 2018

Final version: September 19, 2019

Approved: November 18, 2019