Using Item Response Theory to Measure the Self-Esteem of Workers: a Case Study in a Security and Cleaning Brazilian Company

Usando a Teoria de Resposta ao Item para Medir a Autoestima de Trabalhadores: um Estudo de Caso em uma Empresa Brasileira de Segurança e Limpeza

Waldemar Pacheco Junior[a]; Andréa Cristina Trierweiller[a]; Vera Lúcia Duarte do Valle Pereira[a]; Silvana Ligia Vincenzi[b]; Dalton Francisco de Andrade[a]; Antonio Cezar Bornia[a]; Yuri Borba Vefago[a]

[a]Universidade Federal de Santa Catarina. SC, Brazil.
[b]Universidade Tecnológica Federal do Paraná. PR, Brazil.
*E-mail: andrea.ct@ufsc.br

Abstract

Self-esteem is a personality construct, consisting to the evaluation that the individual makes of himself, attitudes of approval or disapproval shown by verbal reports, and behavior of the individuals, based on Coopersmith (1967, 1989) and Pacheco (2004). The aim of this article is to measure self-esteem among workers of a security and cleaning Brazilian company, using Item Response Theory (IRT). We obtained a sample of 902 workers. The methodological design presents four phases, dimensionality evaluation, items analysis using IRT, creation of scale, and placement of the workers on the scale. We used the Generalized Graded Unfolding Model (GGUM) for establishing a scale with 16 items. We highlighted that eleven items represented a low self-esteem level, assigning the decision-making responsibility to other people or to something external to the will or control of the worker. Five items reflected a high self-esteem. The workers in this scale level showed greater control over their own lives, as a result of their choices and not merely of fate. The self-esteem scale allowed comparing scores between workers, obtaining individual measures. In conclusion, it is possible to apply this scale to other populations with samples of at least one worker, considering the invariance property of IRT.

Keywords: Unfolding Model Psychometric. Item Response Theory. Ergonomics. Scale Development.

Resumo

A autoestima é um construto da personalidade, que consiste na avaliação que o indivíduo faz de si mesmo, nas atitudes de aprovação ou desaprovação demonstradas por relatos verbais e no comportamento dos indivíduos, com base em Coopersmith (1967, 1989) e Pacheco (2004). O objetivo deste artigo é mensurar a autoestima de trabalhadores de uma empresa brasileira de segurança e limpeza, por meio da Teoria de Resposta ao Item (TRI), obtendo-se uma amostra de 902 trabalhadores. O desenho metodológico apresenta quatro fases, avaliação da dimensionalidade, análise dos itens pela TRI, criação da escala e posicionamento dos trabalhadores na escala. Utilizou-se o Generalized Graded Unfolding Model (GGUM) para o estabelecimento de uma escala com 16 itens. Destaca-se que onze itens representaram um baixo nível de autoestima, atribuindo a responsabilidade da tomada de decisão a outras pessoas ou a algo externo à vontade ou controle do trabalhador. Cinco itens refletiram uma elevada autoestima. Os trabalhadores desse nível da escala demonstraram maior controle sobre a própria vida, em função de suas escolhas e não apenas do destino. A escala de autoestima permitiu comparar os escores entre os trabalhadores, obtendo medidas individuais. Em conclusão, é possível aplicar esta escala a outras populações com amostras de pelo menos um trabalhador, considerando a propriedade de invariância da TRI.

Palavras-chave: *Unfolding Model Psychometric. Teoria de Resposta ao Item. Ergonomia. Desenvolvimento de escala.*

1 Introduction

Ergonomics represents a paradigm shift in the relationship between humans and work. Ergonomics criticizes the *Taylorism*, whose application aims to adapt humans to work. *Taylorism* is essentially technocentric: based on the study of performance and productivity, it presupposes the stability of the operator and seeks to demonstrate that the technique, the instruments and the work conditions may be changed. In other words, its first concern is the technique.

Ergonomics defends the opposite principle in the workplace: adaptation of work to individual humans, instead of focusing on the average well-trained and stable worker, according to Wisner (1987). In this sense, the Ergonomics

Regulation Norm (NR-17) (MTE, 1990), according to the Ministry of Labor and Employment of Brazil, considers the promotion of comfort, safety and efficiency to workers as the objective of ergonomic action, thereby aiming to adapt working conditions to the psycho-physiological features of workers.

In a workplace, there usually is a gap between the prescribed task and actual task. In this sense, Guérin *et al.* (2007) asserts that an ergonomist should understand the work to transform it. We must analyze a wide range of factors, including inter and intra-individual variability, thus enabling the introduction of flexible elements into the work organization (ABRAHÃO, 2000). Self-esteem is an aspect of human variability essential in ergonomics studies for understanding the interactions that

occur between individuals and elements of the work system.

This variability is based on assumptions about the differences between subjects (inter-individual) and its singularities. (intra-individual), which, in turn, influence the behavior used to accomplish the work. In ergonomics, the individual does not exist in the abstract; instead, he or she is a human being who thinks, acts and feels, seeking to respond to the imperatives present in the work environment and in human variability. This perspective precludes the prescription and achievement of equal performances across workers (ABRAHÃO, 2000; GUÉRIN *et al.*, 2007; PINHO; ABRAHÃO; FERREIRA, 2003). Furthermore, workers mobilize their cognitive resources to plan and perform work, and many times, those persons responsible for offering prescriptions do not consider the differences between the workers and the unpredictable situations inherent to any work system.

Intra-individual variation affects the way one acts in his or her workplace in physical terms (e.g., postures, gestures, movements, displacements), cognitive terms (e.g., diagnosis, planning, problem solving) and affective terms (experiences of pleasure, suffering, identity construction) (ABRAHÃO, 2000; GUÉRIN *et al.*, 2007; PINHO; ABRAHÃO; FERREIRA, 2003). The presence of variability in a work situation requires workers to invest their intelligence in a constant search for equilibrium. Their knowledge and expertise are directly related to the results of effectiveness, comfort, safety and health at work. In this sense, self-esteem, a variable component of any worker's personality, may be influential (e.g., may affect the *modus operandi,* rhythms of work or regulatory strategies).

The self-esteem construct covers other aspects of personality. Self-esteem is the evaluation (value judgment) that an individual makes and usually maintains about himself or herself. Self-esteem requires an individual to expresses an attitude of approval or disapproval and to indicate the degree to he or she considers oneself to be capable, important and valuable. It is a subjective experience that the individual exposes to others by verbal reports and public expressions of behavior (COOPERSMITH, 1967).

Erol and Orth (2011) examined the development of self-esteem in adolescence and young adults. They found that "at each age, emotionally stable, extraverted and conscientious individual experienced higher self-esteem than emotionally unstable, introverted and less conscientious individuals".

Regarding to self-esteem in the workplace, Stinson *et al.* (2010) showed that, when feedback is consistent with self-esteem, an individual experiences sentiment of authenticity, control and certainty. In contrast, inconsistent feedback generates feelings of discomfort, resulting in efforts to solve the discrepancy between the feedback received and self-esteem.

Pierce *et al.* (1989) presented the concept OBSE — *Organization-Based Self-Esteem*—as the degree to which an individual believes himself or herself to be capable, significant and worthy as a member of an organization. Workers with high levels of OBSE feel like an important part of the organization. A study by Gardner and Pierce (2011) correlated OBSE to satisfaction, motivation and effort expended after receiving negative feedback, among other factors. Organizations can expect that workers with high levels of OBSE are not hostile to negative *feedback*, contrary to the high hostility to negative *feedback* witnessed among workers with false self-esteem (i.e., narcissists).

Regarding self-esteem measurement instruments, Buhrmester, Blanton and Swann Jr (2011) study of implicit self-esteem measurement—via the Implicit Association Test (IAT) and Name-Letter Test (NLT), two of the most common measures of implicit self-esteem concluded that both tests are weak measures of self-esteem, and the literature has not provided consistent support for the validity of either measure. The Rosenberg Self-esteem Scale (RSES) has been widely used and was developed to measure self-esteem in adolescents and was translated into 28 languages (SCHMITT; ALLIK, 2005; ROSENBERG, 1965; ROSENBERG, 1983).

We do not identify instruments for measuring the self-esteem of workers in our literature review. Instead, this article aims to measure self-esteem among workers from a Brazilian company specializing in security and cleaning services using Item Response Theory (IRT), and the article is structured as follows: IRT, Method, Results and Discussion.

2 Material and Method

2.1 Item response theory

IRT is a set of mathematical models that reflects the association between an individual's response to an item and the underlying latent variables using a quantitative measurement scale (EMBRETSON; REISE, 2000; HAMBLETON; SWAMINATHAN; ROGERS, 1991). According to Bortolotti *et al.* (2013, p.2344), IRT permits

> [...] proposing forms of representing the relationship between the probability of a respondent to give a certain response to an item, and its latent trait and items characteristics (parameters) in the knowledge field studied.

IRT appeared to address the limitations of the Classical Test Theory – CTT (LORD; NOVICK, 1968), which considers total scores from the tests. IRT assumes that conclusions do not depend on a test as a whole but on each item, that comprises it.

It is used primarily in education and psychology (DE AYALA, 2009), as well as in other areas, e.g., medicine (DE ROOS; MEARES, 1998; VIDOTTO *et al.*, 2006; DAS; HAMMER, 2005), marketing (BAYLEY, 2001; SINGH, 2004), services (COSTA, 2001), information systems (TEZZA; BORNIA; DE ANDRADE, 2011), genetics (TAVARES; ANDRADE; PEREIRA, 2004), organizational effectiveness (TRIERWEILLER *et al.*, 2012) and environmental management disclosure (TRIERWEILLER *et al.*, 2013).

IRT estimates items and respondent parameters and assumes a model to establish a measurement scale. To estimate these parameters, one can defines a set of variables that cannot be measured directly but that express a theoretical concept in terms of a "construct," an "ability" or a "latent trait," such as the "self-esteem" of workers.

IRT allows comparisons to be made between the latent traits of individuals of different populations when submitted to tests that have some common items. The approach makes possible the comparison of individuals from the same population submitted to totally different tests. This comparison is possible because the IRT has items, rather than the test as a whole, as its central element (BORTOLOTTI *et al.*, 2012).

The evaluation of constructs such as self-esteem is complex due to the subjective nature of the phenomenon. Thereby, the use of IRT allows researchers evaluate this type of phenomenon through the elaboration of a set of items that aims to estimate the latent trait. According to Scherbaum, Finlinson and Tamanini (2006), IRT allows the placement of items and respondents on the same analysis scale.

One feature of IRT (EMBRETSON; REISE, 2000; HAMBLETON; SWAMINATHAN; ROGERS, 1991) is the quantity of information about an item and the test; we may obtain this information by the Item Information Function (IIF) and Test Information Function (TIF). The IIF indicates how much information an item provides on each level of the latent trait, which allows verifying the quality of items, identifying which may be eliminated without compromising the measure of the latent trait (BORTOLOTTI *et al.*, 2012). The TIF represents the sum of all item functions and indicates the quantity of information the test provides at different points along latent trait scale. This function is util because it allows evaluate of whether the instrument measures all levels of the latent trait. One of the most important concepts in psychometrics is the Standard Error (SE) of measurement, and we may estimate it for each level of the latent trait (EMBRETSON; REISE, 2000; HAMBLETON; SWAMINATHAN; ROGERS, 1991; ANDRADE; TAVARES; VALLE, 2000).

There is distinction in mathematical form by function, item feature and number of parameters in the various models of IRT. For example, there are models for binary or polytomous, unidimensional or multidimensional, parametric and nonparametric data. However, based the response process, there are the cumulative and unfolding models (SCHERBAUM; FINLINSON; TAMANINI, 2006). In the cumulative models, the probability of an individual choosing a correct answer increases alongside the value of its latent trait i.e., higher levels of a latent trait translates into higher values of the probability of a correct response, describing a monotonic behavior between the choice of response and latent trait.

Unfolding models are based on processes of response in ideal points, non-monotonic (COOMBS, 1950, 1964; THURSTONE, 1928, 1931). The logics of these models is that the individuals select the option of the response that is closest of his position of the latent trait. Therefore, the probability of an individual provides a response to an item is the function of the distance between the parameters of the individual and the position of the item in the scale (BORTOLOTTI; ANDRADE, 2007). These models likely will have great potential in ergonomics area, since we may describe many answers involving ideal points (ROBERTS; LAUGHLIN; WEDELL, 1999).

Among the cumulative models for items with dichotomous responses or multiple-choice items (corrected as expected/unexpected), the Logistical Models of one, two and three parameters stand out (RASCH, 1960; BIRNBAUM, 1968). We can find application examples of Logistical Model of Two Parameters in Trierweiller *et al.* (2013), and Bortolotti *et al.* (2012).

The models for the polytomous items depend on the nature of the categories of the response. In multiple-choice tests in which category responses are unordered (Nominal Categories Model) (BOCK, 1972). In the cases in which the categories are ordered, the Ordinal Model is preferred, e.g., if we consider the Likert scale in the categories of response, it requires the Graded Response Model (SAMEJIMA, 1969,1997), the Partial Credit Model (PCM) (MASTERS, 1982), the Rating Scale Model (RSM) (ANDRICH, 1978), or the Generalized Partial Credit Model (GPCM) (MURAKI, 1992).

The Graded Response Model (GRM) assumes that the categories of response of an item may be organized among each other. Suppose the scores of the categories of item *i* are arranged in an increasing order and are denoted by k=0; 1; ...; where (m_i +1) is the number of categories of the *i*-th item. The probability of that individual j will pick a particular category or higher than i-th item may require an extension of the Logistical Model with two parameters (ANDRADE; TAVARES; VALLE, 2000):

$$P_{i,k}^{*}(\theta_j) = \frac{1}{1 + e^{-a_i(\theta_j - b_{i,k})}} \qquad (1)$$

where $i = 1, 2, ..., I, j = 1, 2, ..., n$, and $k = 0, 1, ..., m_i$,

where $b_{i,k}$ is the parameter of the difficulty of *k*-th category of the item *i*, and other parameters are analogous to those already defined previously in the model.

In this model, discrimination of a specific response category depends as much as the parameter of inclination, which is common to all response categories, as the distance of adjacent difficulty categories. There is an ordering between the categories' difficulty level for an especific item. According to the classification of its scores, ie,

$$b_{i,1} \leq b_{i,2} \leq ... \leq b_{i,m_i} \qquad (2)$$

The probability of an individual *j* choosing a category *k* at the item *i* is given by the expression:

$$P_{i,k}(\theta_j) = P_{i,k}^*(\theta_j) - P_{i,k+1}^*(\theta_j) = \frac{1}{1+e^{-a_i(\theta_j-b_{i,k})}} - \frac{1}{1+e^{-a_i(\theta_j-b_{i,k+1})}} \quad (3)$$

such that

$$P_{i,0}^*(\theta_j) = 1 \quad \text{e} \quad P_{i,m_i+1}^*(\theta_j) = 0,$$

where a_i is the parameter of item discrimination, which is the same for each category of the item; $b_{i,k}$ is the parameter of *k-th* position from item category *I*; and θ_j represents the level of latent trait.

We verify in a review of the literature about the Unfolding Model that there exists a Hyperbolic Cosine Model (MCH) (ANDRICH; LUO, 1998) and a Generalized Graded Unfolding Model *(GGUM)* (ANDRICH; LUO, 1993). The computational programs for estimating the parameters of these models are GGUM2004 (ROBERTS *et al.*, 2006) and Rummfold (ROBERTS; DONOGHUE; LAUGHLIN, 2000). We represent the GGUM model as follows (ANDRICH; LUO, 1993):

$$P(Z=z|\theta_j)=\frac{\exp\left[\alpha_i\left(z(\theta_j-\delta_i)-\sum_{k=0}^{z}\tau_k\right)\right]+\exp\left[\alpha_i\left((M-z)(\theta_j-\delta_i)-\sum_{k=0}^{z}\tau_k\right)\right]}{\sum_{v=0}^{H}\left[\exp\left(\alpha_i\left[v(\theta_j-\delta_i)-\sum_{k=0}^{v}\tau_k\right]\right)+\exp\left(\alpha_i\left[(M-v)(\theta_j-\delta_i)-\sum_{k=0}^{v}\tau_k\right]\right)\right]} \quad (4)$$

where:

Z_i is an observable response to an item *I*; $z = 0, 1, 2, 3,... H$ corresponds to the level of stronger disagreement; and $z = H$ corresponding to the stronger agreement; and where *H* is the number of response categories observable minus 1; Θ_j is the position parameter of the individual *j* on the latent trait scale (self-esteem); δ_i is the position parameter of the item *i* on the latent trait scale.

α_i is the discrimination parameter of item *i*; π_{ik} is the position parameter of the category threshold of subjective response *k* on the scale of the latent trait, relative to the item's position *i*. Corresponding to the distance value between Θ_j and δ_i, that determines the point at which the *k-th* category of subjective responses suggests the probability of responses above the *(k-1)-th,* category of subjective response for individual *j,* item *i* and τ_{i0} by definition is equal to zero;

M is the number of categories of the subjective responses minus 1 i.e., *M=2H+1*.

In the next section, we present the method used in this article, including a set of items.

2.2 Survey with workers in a Security and Cleaning Brazilian Company

We obtained the data from a survey with a population of 2860 workers in 12 units of a Brazilian company specializing in security and cleaning services. From this population, one can obtain a sample of 902 participants (PACHECO Jr, 2004).

The data collection instrument (Table 1) initially was composed of 20 items covering the following facets of self-esteem: self-respect and self-efficiency (6 items), living consciously (two items), self-acceptance (two items), self-responsibility (two items), self-assertion (two items), living intentionally (two items), personal integrity (two items) and realism (two items). The items presented three categories of responses: Totally Agree (CP), Indifferent (IN) and Totally Disagree (DP).

Table 1 - Items from the Self- Esteem Questionnaire

Item	Description
1	I do not belong to anyone, and no one belongs to me.
2	I trust my thoughts.
3	A defeat is always difficult to overcome.
4	I value and act on my ideas.
5	Success depends on me.
6	I came to the world only to help other people.
7	I like daydreaming to avoid reality.
8	My mistakes are pointless.
9	I am the result of the world that I live in.
10	Things that I do wrong embarrass me.
11	Fate is the owner of peoples' lives.
12	The fulfillment of my desires is my responsibility.
13	You should defend what you think.
14	I trust in myself.
15	The future has no future.
16	Negative things pursue me.
17	I must do what I say.
18	Telling the truth is dangerous.
19	In a relationship, one must always make sacrifices for the other person.
20	Knowing yourself is a heavy much responsibility.

Source: Research data.

The methodological design of this study consisted of four steps: (1) a dimensionality evaluation; (2) an item analysis using IRT; (3) creation of a self-esteem scale; and (4) placement of the workers on the scale.

We performed the evaluation of dimensionality using the factor analysis, and we used the principal components method for extractions of eigenvalues (SPSS, 2007).

We proceeded to analyze items using the cumulative model; we used the Multilog *software* (THISSEN, 1991). However, because the items did not fit into this model, we incorporated an unfolding model with the GGUM 2004 (ROBERTS *et al.*, 2006). *Software.* Finally, we were able to elaborate on the self-esteem scale and the placement of the workers on the scale by virtue of their final scores.

3 Results e Discussion

3.1 Dimensionality

The results of factor analysis indicated that the first factor explained 33.92% (Table 2) of the variance, satisfying the criterion established by Reckase (1979). We found a dominant factor, justifying the use of the unidimensional model of IRT.

Table 2 - Extraction of eigenvalues through principal components analysis

Factor	Eigenvalue	Total Variability	Eigenvalue Cumul.	Cumul. %
1	17.64	33.93	17.64	33.93
2	2.69	5.18	20.33	39.11

Source: Research data.

3.2 Item analysis using the Samejima's Graded Response Model

Table 3 presents the estimates of the item parameters (1, 2, 4, 12, 13, 14, 17 and 19). We eliminated the discrimination parameters (*a*) smaller than 0.5 due their quality unsatisfactory for integrating into the scale.

Table 3 - Estimates of the item parameters according to Samejima's Graded Response Model considering the value of average *b*

Item	Description	*a*	*b1*	*b2*	*b* medium
16	Negative things pursue me.	0.89	-2.56	-0.99	-1.78
8	My mistakes are pointless.	1.13	-1.73	-0.81	-1.27
20	Knowing yourself is a heavy much responsibility.	0.92	-1.67	-0.11	-0.89
18	Telling the truth is dangerous.	0.53	-1.39	0.31	-0.54
5	Success depends on me.	1.16	-0.91	0.45	-0.23
11	Fate is the owner of people's lives.	1.32	-0.74	0.36	-0.19
9	I am the result of the world that I live in.	0.69	-1.16	0.93	-0.12
7	I like daydreaming to avoid reality.	0.94	-1.08	0.88	-0.10
6	I came to the world only to help other people.	0.80	-0.98	1.07	0.04
15	The future has no future.	0.75	-0.40	0.89	0.24
1	I do not belong to anyone, and no one belongs to me.	0.76	-0.56	1.33	0.38
10	Things that I do wrong embarrass me.	0.63	-0.03	1.86	0.92

Source: Research data.

The *b* mean indicates the position of the item on the scale, obtained by averaging the values of *b1* and *b2* for each item. We retained the items, according to the Graded Response Model, evaluated respondents with low self-esteem alone; hence, the items' position orders did not coherent to the nature of the latent trait (viz., self-esteem) and the items' content. The results of this model were unsuitable with regarding to the latent trait, then, we investigated the unfolding model with the objective of defining the scale.

3.3 Analysis of the items using the Unfolding Method (GGUM)

We chose to use an unfolding model because the resulting values of the item parameters were more consistent and coherent with a latent trait nature (self-esteem) than were the results based on Samejima's Graded Response Model. We can verify that the latent trait of self-esteem is not cumulative, in contrast to the quality-of-life latent trait described by Lin (2007), which exhibited good item discrimination by means of Samejima's Graded Response Model.

We estimated the items parameters through the GGUM model of Roberts, Donoghue and Laughlin (2000), and we excluded four items (viz., 1, 17, 18 and 19), respectively: "I do not belong to anyone, and no one belongs to me". "I must do what I say." "Telling the truth is dangerous", and "In a relationship, one must always make sacrifices for the other person". We made these exclusions due the *infit* and *outfit* values did not conform to the parameters that would be expected according to Karabatsos (2000), Smith, Schumacker and Bush (1998), Engelhard Jr (1992) and Lunz, Wright and Linacre (1990). Moreover, the discrimination parameter α for each of these items was less than 0.5, indicating that these items do not appropriately discriminate among the various levels of the latent trait.

We calibrated and organized the parameters of the 16 items according to delta (δ) which represents an item's position on the scale (Table 4). The table gives estimates for the parameters of the items (e.g., δ and π), the discrimination parameter of each item (α) and respective statistics of adjustments (i.e., *infit* and *outfit*).

Table 4 - Estimates for the Items' Parameters

Items Description	*d*	*a*	*t₁*	*t₂*	*Infit*	*Outfit*
6. I came to the world only to help other people.	-0.9	0.7	-1.2	-0.5	0.9	0.9
3. A defeat is always difficult to overcome.	-0.8	0.7	-0.6	-0.4	1.0	0.9
10. Things that I do wrong embarrass me.	-0.7	0.7	0.1	-0.5	0.9	0.9
9. I am the result of the world that I live in.	-0.7	0.6	-0.9	-1.2	1.0	1.0
15. The future has no future.	-0.7	0.7	0.0	-1.2	1.0	0.9
5. Success depends on me.	-0.6	1.0	-1.1	-0.9	1.0	1.0
11. Fate is the owner of peoples' lives.	-0.6	1.1	-1.0	-1.0	1.0	1.0
7. I like daydreaming to avoid reality.	-0.6	0.9	-1.3	-0.4	1.0	1.0
8. My mistakes are pointless.	-0.4	1.0	-1.1	-2.1	1.1	1.1
20. Knowing yourself is a heavy much responsibility.	-0.4	0.9	-1.3	-1.3	1.1	1.1
16. Negative things pursue me.	-0.3	0.9	-1.7	-2.0	1.1	1.2
4. I value and act on my ideas.	0.5	1.0	-3.5	-2.6	1.2	1.2
13. You should defend what you think.	0.6	1.0	-3.5	-2.6	1.1	1.0
2. I trust my thoughts.	0.6	1.0	-3.4	-2.9	1.2	1.1
12. The fulfillment of my desires is my responsibility.	0.7	0.7	-2.9	-3.8	1.2	1.1
14. I trust in myself.	1.2	1.1	-2.9	-3.6	1.1	0.7

Source: Research data.

We list items in ascending order according to the value of the parameter *delta* (δ), ranged from -0.9 to 1.2, which corresponds to the content and nature of the latent trait (viz. self-esteem). For example, individuals who presented the lowest levels of self-esteem answered the item 3, "A defeat is always difficult to overcome" (δ= -0.8), thereby reflecting how this model is based on proximities between items and individuals. The parameter *alpha* (α) describes the discrimination parameter for the item; in the case of item 3 (*α*=0.7), the item adequately discriminates based on its content.

Tau is the parameter of the threshold position of the subjective response. This value represents the position of the subjective limits between the options of response (for instance, "disagree totally" and "indifferent") in relation to the position parameter of the item (δ). The value of *tau*-1 for item 3, "A defeat is always difficult to overcome," indicates that the category of subjective response to agree is located in the range -1.2< δ < -0.4. The *infits* and *outfits* of each item had values close to 1, suggesting that the items fit into the model.

Figure 1 presents the Test Information Function given by the sum of the information from all the items. This curve demonstrates the place where there is greater precision for estimation of the self-esteem of workers, and hence, more information exists for the items that discriminate low self-esteem.

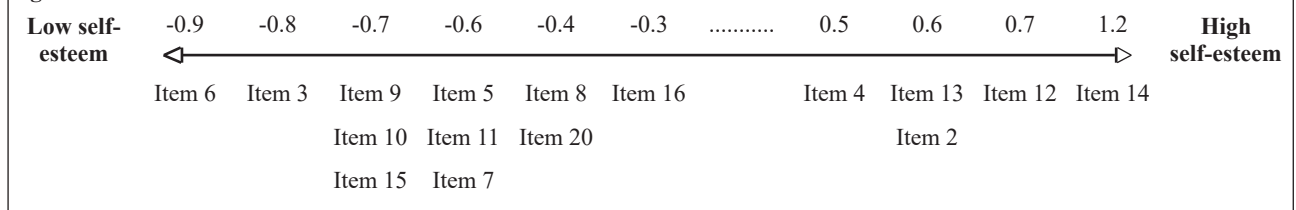**Figure 1 -** Test Information Function (TIF) of self-esteem
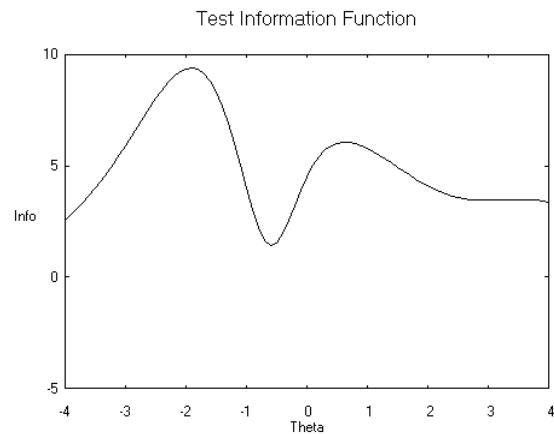
Source: Research data.

The information curve of this measure instrument of self-esteem assures that the items are suitable for the region of interest, reflecting the quality of the items, according to Baker and Kim (2004).

3.4 Creation of the self-esteem scale

We calibrated the data were from the GGUM model considering the latent trait, self-esteem and its possible behavior. We identified two regions on the scale (Figure 2 shows the positions of the items on the scale of self-esteem):

Figure 2 - The self-esteem scale derived from the GGUM model

Low self-esteem	-0.9	-0.8	-0.7	-0.6	-0.4	-0.3	....	0.5	0.6	0.7	1.2	High self-esteem
	◁										▷	
	Item 6	Item 3	Item 9	Item 5	Item 8	Item 16		Item 4	Item 13	Item 12	Item 14	
			Item 10	Item 11	Item 20				Item 2			
			Item 15	Item 7								

Source: Research data.

The parameter δ is the base of the scale construction, which varied between -0.9 and 1.2. We organized the levels of the scale as follows: low self-esteem (-0.9 to -0.3) and high self-esteem (0.5 to 1.2). Therefore, the two levels of self-esteem obtained from the model evidenced a clear division.

The items 6, 3, 10, 9. 15, 5, 11, 7, 8, 20 and 16 represent a low self-esteem level, and the items 4, 13, 2, 12 and 14, high self-esteem. The transition between the scale levels was gradual for the GGUM model and did not always allow a precise delimitation at a cutoff (DEMARS; ERWIN, 2003). Nevertheless, in this survey, the self-esteem levels showed an obvious delimitation.

We can evaluate the self-esteem by levels: low, average and high (ROSENBERG, 1965; ROSENBERG, 1983). Feelings of inability, inadequacy regarding life and incapacity to overcome challenges characterize low self-esteem. High self-esteem reflects feelings of confidence and competence.

Average self-esteem floats between such feelings of adequacy and inadequacy, manifesting itself in inconsistency of behavior.

We obtained that the items 6, 3, 10, 9. 15, 5, 11, 7, 8, 20 and 16 represent a low self-esteem level, as follows: "I came to the world only to help other people." "A defeat is always difficult to overcome." "Things that I do wrong embarrass me." "I am the result of the world that I live in." "The future has no future." "Success depends on me." "Fate is the owner of peoples' lives." "I like daydreaming to avoid reality." "My mistakes are pointless." "Knowing yourself is a heavy much responsibility." and "Negative things pursue me." These items assign the decision-making responsibility to other people or to something external to the will or control of the worker.

Items 4, 13, 2, 12 and 14 reflect a high self-esteem: "I value and act on my ideas." "You should defend what you think." "I trust on my thoughts." "The fulfillment of my desires is my responsibility." and "I trust in myself." The workers in this scale level show greater control over their own lives. They

assume that their lives are the result of their choices and not of fate.

The final scale included more items focused on the measurement of low self-esteem. According to Kolen and Brennan (1995). IRT permits the addition of new items by increasing the set of items using the equalization method. We can include items, since using psychometric procedures that support the validity and reliability of the instrument (NUNNALLY, 1978).

In this article, we can verify the advantages of the IRT model that are mentioned by Hambleton and Swaminathan (1985) and Embretson and Reise (2000) because the order of items generated from our procedure allowed us to localize the items according to their content and the position of the respondents with respect to their self-esteem levels.

The self-esteem scale allowed comparing scores between workers, obtaining individual measures. It is possible to apply this scale to other populations with samples of at least one worker, considering the invariance property of IRT (EMBRETSON; REISE, 2000; HAMBLETON; SWAMINATHAN; ROGERS, 1991).

We identified four items that presented low discrimination, and missing adjustment, which require content modification. The analysis of the data used to develop this instrument differs from instruments derived from Classical Test Theory (CTT) based on the scoring. In CTT, the evaluation of the same latent trait through a set of items with varied levels of difficulty produce different scores - i.e., there is dependence on the set of items that we used (HAMBLETON; SWAMINATHAN; ROGERS, 1991). According to Bortolotti *et al.* (2012), IRT was developed mainly to surpass the limitations of CTT. The main feature of this theory is the focus on considering each item individually such that the conclusions do not depend on the instrument as a whole but on each item that comprises it.

From the proposal instrument, we identified some attitudes that revealead the worker self-esteem, related to variability aspects, which is an important condition for ergonomic action (FERREIRA; MENDES, 2003). Although the scale can contribute to diagnostic assessment of self-esteem, there is no possibility of finding, at work environment, elements in favor to the construction of personality. Nevertheless, the constraints of work organization may lead to personality changes, contributing negatively to the health of workers (GUÉRIN *et al.*, 2007).

Moreover, self-esteem is a reflex of learning and individual history manifested in personality that depends on both internal and external factors. According to Coopersmith (1989, p.2): "[...] Children are not born with concerns of being good or bad, smart or stupid, lovable or not. They develop these ideas. They make self-images... strongly based at the way they were treated by meaningful people [...]".

We used the unfolding model to the criation of the scale, following the order of the latent trait's nature, providing the measurement of self-esteem in the workers. We generated

a scale with levels of high and low self-esteem, confirming findings of Roberts, Donoghue and Laughlin (2000), and Demars and Erwin (2003), who has found a fit of the latent trait to the model.

3.5 Placement of workers on the self-esteem scale

Figure 3 shows the frequency distribution of self-esteem reported by the workers using the scale generated from the model.

Figure 3 - Distribution of workers with respect to their levels of self-esteem
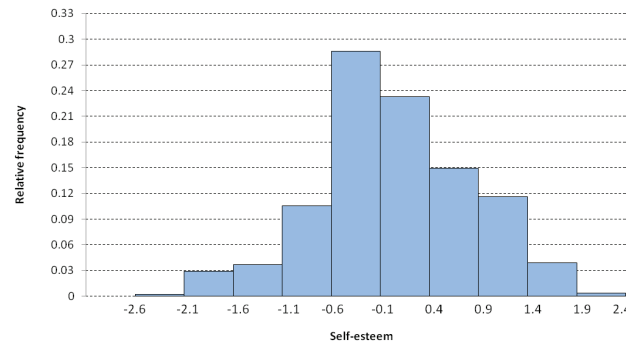
Source: Research data.

Approximately 63% of workers exhibited low self-esteem, whereas 37% exhibited high self-esteem (level > 0.5). However, the self-esteem construct covers other aspects of personality (MRUK, 2006), which may to suffer damage in constraint situations on work (GUÉRIN *et al.*, 2007).

The relationship between work and health is not influenced solely by work demands because workers engage in activities based on own their personalities and histories. Confrontation between the personal feature of workers and operating strategies will unfold positively or negatively with respect to their health (GUÉRIN *et al.*, 2007).

4 Conclusion

Based on the unfolding model, we can infer that the probability of a worker responding to an item depends on the distance (i.e., absolute difference) between the worker's position and the item's position on the scale. The probability of the worker is agreeing with (i.e., giving a positive response to) the item increases with decreasing distance between the individual's position and the position's item on the self-esteem latent trait's scale. I.e., we can apply the unfolding model when respondents indicate their agreement with a set of items situated on a bipolar scale in this case, high and low self-esteem.

The final instrument included 16 items that showed adequate fit and discrimination, and we placed the items on the scale according to their content (low and high self-esteem). Moreover, we placed the workers on the same scale according to their self-esteem level.

According to results, most workers presented low self-esteem, which may be related to constraints of their

organization, compromising to their health. Although, it is not an ergonomics objective to act directly between the self-esteem and the organization, however, some situations of work show pathologies unsuitable by its material features. Thus, the ergonomist should act over the situation and work relations attentive to operating strategies of workers, and collaborating to new solutions regarding to difficulties found, and conflicts that may affect the self-esteem level.

The article focus is not on the process of influencing workers' self-esteem but on how the attitudes may represent self-esteem, measured from our instrument. For example, the scale reflects high self-esteem attitudes, as shown in items 2, 12, 14, respectively: "I trust my thoughts." "The fulfillment of my desires is my responsibility," "I trust myself" – reflect the right side of the scale, whereas items 10, 3 and 6 ("What I do wrong embarrasses me," "A defeat is always hard to overcome," and "I came to the world only to help other people") reveal low self-esteem, the left side on the scale.

Based on our results, we showed the viability of measuring the self-esteem of workers through IRT, which allows evaluating the quality and property of every item individually, placing items and respondents on the same scale. This approach is particularly useful for the evaluation of constructs that are difficult to observe or are abstract, such as self-esteem.

In the future research, we suggest the elaboration of the items to cover medium and high self-esteem. According to Mruk (2006) a good self-esteem instrument should consider several aspects of phenomenon and be capable of responding to particular situations in an individual's life regarding the self-steem.

This paradigm shifts in the construction and validation of measuring instruments offered by IRT creates the possibility of new perspectives in evaluating other constructs of ergonomics, searching the understanding the man as central element and active in the process of work analysis. Self-esteem is an aspect relevant to ergonomics interventions because knowing a workers' attitude may contribute to understanding the interactions, he or she has with the elements of the work system.

Acknowledgments

References

ABRAHÃO, J.I. Reestruturação produtiva e variabilidade do trabalho: uma abordagem da ergonomia. *Psicol. Teor. Pesq.*, v.16, n.1, p.49-54, 2000. doi:10.1590/S0102-37722000000100007

ANDRADE, D.F.; TAVARES, H.R.; VALLE, R.C. *Teoria de resposta ao item*: conceitos e aplicações. Associação Brasileira de Estatística - ABE, 4º SINAPE, 2000.

ANDRICH, D. A rating formulation for ordered response categories. *Psychmetrika*, v.43, p.561-573, 1978.

ANDRICH, D.; LUO, G. A hyperbolic cosine latent trait model for unfolding dichotomous single-stimulus responses. *Applied Psychological Measurement*, v.17, p.253-276, 1993. doi: 10.1177%2F014662169301700307

ANDRICH, D.; LUO, G. Rummfoldtm for WindowsTM: a program for unfolding pairwise preferences, Computer Program. *Murdoch, Western Australia: Social Measurement Laboratory, Murdoch University*, 1998.

BAKER, F. B.; KIM, S. *Item response theory*: parameter estimation techniques. New York: Marcel Dekker, 2004.

BAYLEY, S. Measuring customer satisfaction. *Eval. J.Australasia*, v. 1, n. 1, p. 8-17, 2001.

BIRNBAUM, A. Some latent trait models and their use in inferring an examinee's ability. In: LORD, F. M., NOVICK, M. R., *Statistical theories of mental test scores*, Reading-MA: Addison–Wesley, 1968.

BOCK, R. D. Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika,* v. 37, n. 1, p. 29-51, 1972.

BORTOLOTTI, S. L. V; ANDRADE, D. F. Aplicação de um modelo de desdobramento graduado generalizado-GGUM da teoria da resposta ao item. *Estud. Aval. Educ.*, v.8, n.37, p.157-188, 2007. doi:10.18222/eae183720072094

BORTOLOTTI, S.L.V. et al. Avaliação do nível de satisfação de alunos de uma instituição de ensino superior: uma aplicação da teoria da resposta ao item (Consumer satisfaction and item response theory: creating a measurement scale.). *Gestão & Produçã*, v.19, n.2, p.287-302, 2012.

BORTOLOTTI, S.L.V. et al. Relevance and advantages of using the item response theory. *Quality & Quantity*, v. 47, p. 2341-2360, 2013. doi:10.1007/s11135-012-9684-5

BUHRMESTER, M.D.; BLANTON, H.; SWANN JUNOR, W.B. Implicit Self-Esteem: Nature, Measurement, and a New Way Forward. *J. Personal. Soc. Psychol.,* v.100, n.2, p.365-385, 2011. doi: 10.1037/a0021341

COOMBS, C.H. Psychological scaling without a unit of measurement. *Psychol. Rev*, v.57, p.145-158, 1950. doi:10.1037/h0060984

COOMBS, C. H. *A theory of data*. New York: Wiley, 1964.

COOPERSMITH, S. *The antecedents of self-esteam*. San Francisco, 1967.

COOPERSMITH, S. *Coopersmith self-esteem inventory*. Palo Alto-CA: Consulting Psychologists Press, 1989.

COSTA, M. B. F. *Technique derived from Item Response Theory (IRT) applied to the service sector*. Master of Science, Graduate Program in Numerical Methods in Engineering, Universidade Federal do Paraná, Curitiba-PR, Brasil, 2001.

DAS, J.; HAMMER, J. Which doctor? Combining vignettes and item response to measure clinical competence. *J. Develop. Econ.s*, v.78, p.348-383, 2005. doi: 10.1016/j.jdeveco.2004.11.004

DE AYALA, R. J. *The theory and practice of item response theory*. New York: Guilford Press, 2009.

DEMARS, C. E.; ERWIN, D. Revising the scale of intellectual development: application of an unfolding model. *J. Coll. Student Develop.*, v.44, n.2, p.1-17, 2003.

DE ROOS, Y.; MEARES, P. A. Aplication of Rasch analysis: exploring differences in depression between African American

and White children. *J. Soc. Serv. Res.*, v.23, n.3/4, p.93-107, 1998. doi:10.1300/J079v23n03_06

EMBRETSON S.; REISE, S.P. *Item response theory for psychologists*. New Jersey: Lawrence Erlbaum Associates, 2000.

ENGELHARD JUNOR, G. The measurement of writing ability with a many-faceted Rasch model. *Appl. Measur. Educ.*, v.5, p.171-191, 1992. doi:10.1207/s15324818ame0503_1

EROL, R. Y.; ORTH, U. Self-Esteem development from age 14 to 30 years: a longitudinal Study. *J. Personal. Soc. Psychol.*, v.101, n.3, p.607-619, 2011. doi:10.1037/a0024299

FERREIRA, M.C.; MENDES, A.M. *Work and risks of illness*: the case of tax auditors from Brazilian Social Security. Brasília: Ler, Pensar, Agir, 2003.

GARDNER, D. G.; PIERCE, J. L. A question of false self-esteem Organization-based self-esteem and narcissism in organizational contexts. *Journal of Management Psychology*, v. 26, n. 8, p. 682-699, 2011. doi:10.1108/02683941111181770

GUÉRIN, F.; LAVILLE, A.; DANIELLOU, F.; DURAFFOURG, J.; KERGUELEN, A. *Understanding and transforming work*: the practice of ergonomics. Lyon: ANACT, 2007.

HAMBLETON, R. K.; SWAMINATHAN, H. *Item Response Theory Principles and Applications*. Boston: Kluwer Academic Publishers, 1985.

HAMBLETON, R. K.; SWAMINATHAN, H.; ROGERS, H. J. *Fundamentals of item response theory*. Newbury Park-CA: Sage, 1991.

KARABATSOS, G. A critique of Rasch residual fit statistics. *J. Appl. Measurement*, v.1, n.2, p.152-176, 2000.

KOLEN, M. J.; BRENNAN, R. L. *Test Equating*. New York: Spring, 1995.

LIN, T. H. Identifying optimal items in quality of life assessment. *Quality & Quantity*, v. 41, p. 661–672, 2007. doi:10.1007/s11135-006-9017-7

LUNZ, M. E.; WRIGHT, B. D.; LINACRE, J. M. Measuring the impact of judge severity on examination scores. *Appl. Measur. Educ.*, v.3, p.331-345, 1990. doi:10.1207/s15324818ame0304_3

LORD, F. M.; NOVICK, M. R. *Statistical theories of mental test score*. Reading: Addison-Wesley, 1968.

MASTERS, G. N. A Rasch model for partial credit scoring. *Psychometrika*, v.47, p.149–174, 1982. doi:10.1007/BF02296272

MOREIRA, A. L. Ergonomia: mitos, verdades e controvérsias. *Revista de Ciências Gerenciais*, v. 9, n. 11, p. 54-61, 2005. doi:10.17921/1415-6571.2005v9n11p54-61

MRUK, C. J. *Self-esteem research, theory, and practice*: Toward a positive psychology of self-esteem (3rd ed.). New York: Springer, 2006.

MTE – Ministério do Trabalho e Emprego. *Norma Regulamentadora. N° 17 (NR17) – Ergonomia*. (Ministry of Labor and Employment (1990). Ergonomics Regulation Norm No. 17), 1990.

MURAKI, E. A generalized partial credit model: application of an EM algorithm**.** *Applied Psychological Measurement*, v. 16, n. 2, p. 159-176, 1992. doi:10.1002/j.2333-8504.1992.tb01436.x

NUNNALLY, J. *Psychometric theory*. New York: McGraw-Hill, 1978.

PACHECO Jr, W. *Contingency approach in Human Resources Management*. 2004. Thesis (Production Engineering Post Program), Florianópolis, Federal University of Santa Catarina, Brazil, 2004.

PIERCE, J.; GARDNER, D. G.; GUMMINGS, L.; DUNHAM, R. B. Organizational-based self-esteem: contruct definition, measurement, and validation. *The Academy of Management Journal*, v. 32, n. 3, 1989.

PINHO, D. L. M.; ABRAHÃO, J. I.; FERREIRA, M. C. As estratégias operatórias e a gestão da informação no trabalho de enfermagem, no contexto hospitalar. *Revista Latino-Americana de Enfermagem*, v. 11, n. 2, p. 168–176, 2003. doi:10.1590/S0104-11692003000200005

RASCH, G. *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research, 1960.

RECKASE, M. D. Unifactor latent trait models applied to multifactor tests: Results and implications. *Journal of Educational and Behavioral Statistics*, v. 4, p. 207-230, 1979. doi:10.3102/10769986004003207

ROBERTS, J. S; LAUGHLIN, J. E.; WEDELL, D. H. Validity issues in the Likert and Thurstone approaches to attitude measurement. *Educational and Psychological Measurement*, v. 59, p. 211–233, 1999. doi:10.1177/00131649921969811

ROBERTS, J. S.; DONOGHUE, J. R.; LAUGHLIN, J. E. A general model for unfolding Unidimensional polytomous responses using item response theory. *Applied Psychological Measurement*, v. 24, n. 1, p. 3–32, 2000.

ROBERTS, J. S. *et al.* A Windows-based Program to Estimate Parameters of the Generalized Graded Unfolding Model. *Applied Psychological Measurement*, v. 30, n. 1, p. 64-65, 2006. doi:10.1177/0146621605280141

ROSENBERG, M. *Society and the adolescent self-image*. Princeton-NJ: Princeton University Press, 1965.

ROSENBERG, M. *La autoestima del adolescente y la sociedad*. Buenos Aires: Paidós, 1983.

SAMEJIMA, F. An estimation of latent ability using a response pattern of graded scores. *Psychometric Monograph*, 1969.

SAMEJIMA, F. Graded response model. In: VAN DER LINDEN, W. J.; Hambleton, R. K. (Eds.). *Handbook of modern item response theory*, Springer, New York, p.85–100, 1997.

SCHERBAUM, C. S.; FINLINSON, S.; TAMANINI, K. Applications of item response theory to measurement issues in leadership research. *The Leadership Quarterly*, v.17, n.4, p.366-386, 2006. doi:10.1016/j.leaqua.2006.04.005

SCHMITT, D.P.; ALLIK, J. Simultaneous administration of the Rosenberg Self-Esteem Scale in 53 nations: Exploring the universal and culture-specific features of global self-esteem. *Journal of Personality and Social Psychology*, v. 89, p. 623–642, 2005. doi:10.1037/0022-3514.89.4.623

SINGH, J. Tackling measurement problems with item response theory: principles, characteristics, and assessment, with an illustrative example. *Journal of Business Research*, v. 57, p. 184-208, 2004. doi:10.1016/S0148-2963(01)00302-2

SMITH, M.; SCHUMACKER, R. E.; BUSH, M. J. Using item mean squares to evaluate fit to the Rasch model. *Journal of Outcome Measurement*, v. 2, p. 66–78, 1998.

SPSS. Statistical Package for the Social Sciences Inc. *Released 2007*. SPSS for Windows, Version 16.0. Chicago, SPSS Inc, 2007.

STINSON, D. A. *et al.* The regulatory function of self-esteem: Testing the epistemic and acceptance signaling systems. *Journal of Personality and Social Psychology*, v. 99, n. 6, p. 993–1013, 2010.

TAVARES, H. R; ANDRADE, D. F.; PEREIRA, C. A Detection of determinant genes and diagnostic via item response theory. *Genetics and Molecular Biology*, v.27, n.4, p.679-685, 2004. doi:10.1590/S1415-47572004000400033

TEZZA, R.; BORNIA, A. C.; DE ANDRADE, D. F. Measuring web usability using item response theory: principles, features and opportunities. *Interacting with Computers*, v. 23, p. 167-175, 2011. doi:10.1016/j.intcom.2011.02.004

THURSTONE, L. L. A law of comparative judgments. *Psychological Review*, v.34, p.278–286, 1928.

THURSTONE, L. L. The measurement of social attitudes. *Journal of Abnormal Psychology*, v. 26, n. 3, p. 249–269, 1931. doi:10.1037/h0070363

THISSEN, D. *MULTILOG user's guide*: multiple categorical item analysis and test scoring using item response theory. Scientific Software Int., Chicago, 1991.

TRIERWEILLER, A. C. *et al.* Measuring organizational effectiveness in information and communication technology companies using item response theory. *Work (Reading, MA)*, v. 41, p. 2795-2802, 2012. doi:10.3233/WOR-2012-0526-2795

TRIERWEILLER, A. C.; *et al.* Measuring Environmental Management Disclosure in Industries in Brazil with Item Response Theory. *Journal of Cleaner Production*, v. 47, p. 298-305, 2013. doi:10.1016/j.jclepro.2012.10.025

VIDOTTO, G. *et al.* A new questionnaire specifically designed for patients affected by chronic obstructive pulmonary disease: the Italian Health Status Questionnaire. *Respiratory Medicine*, v. 100, n. 5, p. 862-870, 2006. doi:10.1016/j.rmed.2005.08.024

WISNER, A. *Por dentro do trabalho*: ergonomia, método e técnica. São Paulo: FTD/Oboré, 1987.