

## TO IMPROVE THE EMPLOYEE ASSESSMENT PROCEDURES – DEVELOPMENT OF RATIO WITH THE USE OF IRT MODELS

**Anna Zięba (ORCID: 0000-0002-0575-8318)**

Faculty of Economic Sciences  
Wrocław University of Economics, Poland  
e-mail: anna.zieba@ue.wroc.pl

**Abstract:** Taking into account the trends in selecting and evaluating employees connected with the use of more and more objective tools, as well as more and more crucial human resources issues in an enterprise, it seems necessary to develop more and more innovative methods and tools. The objective of this paper is to develop a ratio to assess the employee on the basis of several traits at the same time as well as the achievement of specific targets in the company growth. Such ratios are rarely discussed in the literature, however, the analysis can provide a lot of useful information about a specific issue.

**Keywords:** GRM, IRT models, ratio, employee assessment, periodic assessment

### INTRODUCTION

The thesis that the employee is the main factor determining the competitiveness and the human capital is one of the most crucial resources of the company has already become a canon in the theories of human resources management and organization psychology. The activities supporting the employee development in the scope of formal as well as interpersonal qualifications are obvious yet often underappreciated. The organizations which realize the value provided by the motivated and committed employee tend to “diagnose” their possibilities and weaknesses as precisely as possible, and the decisions regarding their career paths are corrected regularly on the basis of various kinds of employee assessments [Verbruggen 2010]. Those employee assessments are made for the purpose of promotion, career path development, trainings or in connection with the company reorganization. Such assessments also provide useful information.

Knowledge and information are the two factors which at present determine the life of the organization. That is why the problem which often arises is how to make assessments so that they are an effective tool in the organization management process.

A lot of ratios, such as absenteeism, productivity, efficiency are used in the employee assessment process. All of the ratios are important, however, it should be considered how to distinguish one employee from another on the basis of those ratios. Is it enough to say that their assessment score is the same (e.g. in regard of efficiency) which would presumably mean that they are equally "good" and they should get the same bonus? The most common system relies on that very assumption. However, the employee's attitude to work (their motivation), commitment to all kinds of additional activities, cooperation with other workers and the general work discipline are also worth evaluating. So in effect a problem emerges – namely what tool should be applied to be able to assess the employees, comparing several traits at the same time (or evaluating the execution of several targets at the same time).

So far a lot of attention has been paid in the literature to the issue of employee assessment. Different assessment methods have been discussed [Sidor-Rządkowska 2000, Jędrzejczak 2000]. The authors agree that it is difficult to indicate a solution which would replace the employee assessment and that is why the question which is asked now in the subject literature is not "why assess?" but "how to assess?" What is stressed is the need to use assessments in the process of developing the company growth strategies, in controlling the achievement of set targets [Juchnowicz 2003]. However, it is known that in spite of the growing interest in the problem of employee assessment little attention has been paid so far to that issue from the perspective of the impact of latent traits on the assessment results. The most frequent ratios suggested in the literature do not include the differences between employees caused by different level of intensity of latent traits.

That is why the objective of this paper is to develop a ratio which would support the decision making process and could be employed to assess the employee in respect of several traits at the same time as well as the achievement of specific targets of the company development. It is important then to develop the measuring tools to formulate correct conclusions, affect and facilitate decision making.

Based on the issues mentioned above the following hypothesis was put forward:

Hypothesis: a ratio developed with the use of IRT models is a tool that can be applied to assess the employee in respect of several traits at the same time and the achievement of specific targets of the company development.

## EMPLOYEE ASSESSMENT

Employee assessment is a process to evaluate personal traits, attitudes, behaviors and the assigned task completion level. The assessment results are the

basis for planning trainings and employee development as well as remuneration schemes, making decisions on promotions, awards, pay rises or dismissals.

As a result of employee assessment it is possible to provide their employers and the employees themselves with information on the results, on how they are perceived by other employees, on the possibilities for growth in a given position.

It is important why assessments are made as well as what objectives of the assessment are. The objectives of periodic assessments should be then the starting point [Listwa 1999, Szałkowski 2000].

The assessment criteria should correspond to the question: What are we going to assess? All applicable assessment criteria are usually divided into four groups [Ludwiczynski 2014]: qualification criteria, effectiveness criteria, behavioral criteria and personality criteria. Most controversy is caused by the personality criteria which regard the traits of a given employee which determine their behavior and attitude at work. Some examples of personality criteria used in practice include responsibility, creativity, assertiveness, resistance to stress. Some also mention intelligence, talents and temperament [Pawlak 2003, Poczowski 2003]. The discussions over that criterion regard the issue of dependence between human personality and work results.

A reliable and accurate employee assessment is difficult. What are the problems connected with the employee assessment process? The easiest thing to assess is the employee effectiveness on the basis of the results e.g. sales performance. It is more difficult to make assessment on the basis of behaviors. Most organizations apply the following ratios: employee productivity (the ratio of the number of manufactured products to the number of hours worked by the employees), completion of the production plan (the number of manufactured finished products on a shift, whole day in comparison to the number assumed in the plan), absenteeism (absences in hours or days in comparison to the whole amount of time in hours or days.) These are only examples because the type and number of the ratios should correspond to the company needs.

To sum up the above discussion it is worth repeating the question asked in the introduction: How to distinguish one employee from another? Is it enough to say that their assessment score is the same, which would mean that they are equally "good?" The following sections of the paper explain the claim that the influence of other traits which are not directly observable on the results should be taken into account and the results should be differentiated.

## GRM – KEY INFORMATION

The most obvious reason for the development of multi-category models of responses is the fact that multi-category questions are most often used in various kinds of studies. Depending on whether the categories are ordinal or not, there are various types of models.

Samejima initially proposed Graded Response Model used for the analysis of multi-category questions. Each question no.  $j$  in this model (GRM) is characterized by two kinds of parameters: parameter  $\beta_j$  describing the item discrimination parameter and by parameters known as item location thresholds  $\alpha_{jm}$ , where  $m=1,2,\dots,M$  means the number of categories [Samejima 1997]. For instance: in the case of a question with three categories of responses there are two threshold values:  $\alpha_{j2}$  - threshold separating the first from the second category and threshold  $\alpha_{j3}$  - threshold separating the second from the third category. Parameters  $\alpha_{jm}$  indicate the latent trait level that is necessary to provide a response above that threshold value.

Each question no.  $j$  includes  $K_j$  possible response categories. The respondent chooses one of the categories (the possibility of choosing several categories within one question is a different issue). The probability of response to question  $j$  can be defined in each of the categories for person  $i$  and question  $j$  with  $K_j$  response categories. That probability is designated as  $\pi_{ijk}$ ,  $k = 0, \dots, K_j - 1$ . These probabilities within each question sum up to 1.

The Samejima's model is based on the accumulated probability. The function describing the probability of providing response to question  $j$  in category  $k$  was defined as follows:

$$\log\left(\frac{\pi_{ijk}}{1 - \pi_{ijk}}\right) = \theta_i - \alpha_{jk} \quad (1)$$

$\theta_i$  - parameter related to respondent  $i$ , indicating the degree of intensity of the analyzed latent trait.

The probability of choosing the  $k$  or higher category and the  $(k+1)$  or higher category, in the case of question  $j$ , is defined as follows (Samejima 1997):

$$P(X_{ij} \geq k) = \frac{\exp((\theta_i - \alpha_{jk})\beta_j)}{1 + \exp((\theta_i - \alpha_{jk})\beta_j)} \quad (2)$$

and

$$P(X_{ij} \geq k + 1) = \frac{\exp((\theta_i - \alpha_{j(k+1)})\beta_j)}{1 + \exp((\theta_i - \alpha_{j(k+1)})\beta_j)} \quad (3)$$

In order to avoid the situation in which

$$P(X_{ij} = k) = P(X_{ij} \geq k) - P(X_{ij} \geq k + 1) < 0 \text{ it is assumed that } \alpha_{jk} < \alpha_{j(k+1)}.$$

## DEVELOPMENT OF THE RATIO – RESEARCH METHOD

The assessment ratios are an important element affecting the implementation of a strategy assumed in an organization. With the use of the ratios it is easier to systematically monitor the completion of the objectives and take actions eliminating the inadmissible (too big) deviations of the ratios from the assumed target values (norms.)

In the case of employee assessment it is necessary to analyze many different situations and many different traits. The assessment is made by comparing the traits, qualifications, behaviors of one employee towards other employees or against a set standard. In effect then the assessment should have some point of reference [Ludwiczynski 2014]. It is difficult, however, to compare employees and at the same time analyze all studied traits. What is needed then is the ratios with which it would be possible to evaluate at the same time the completion of several goals by the employees.

The ratio suggested here should be calculated on the basis of selected traits (most crucial for a given position from the point of view of the employer.) The ratio shall be developed in several stages.

### **Stage I**

All studied traits should be comparable. The most common approach is to change all indicators into stimulants. If, however, the analyzed values are expressed in different units of measure, they should be standardized.

### **Stage II**

The level of intensity of the studied traits is estimated with the use of the latent trait models. The directly unobservable traits (latent traits) are measured with the useful tool called IRT models (*Item Response Theory*.) With the use of IRT models it is possible to evaluate the relationship between the responses to the questions and the level of intensity of the analyzed traits (see e.g. [Ayala 2009, Wilson 2004]). One of the features of those models is the use of observable behaviors to estimate the level of intensity of the latent trait which is studied.

### **Stage III**

The value which is the most desirable from the point of view of the observer is identified from among all values of the selected trait. Usually this is the maximum value from among all values of a given trait.

### **Stage IV**

Most often the similarity of observations is determined with the use of the distances between the observations. Large distances mean a small probability and the other way around.

The most common methods of determining the distances are based on the following metrics: the Minkowski distance, the Czebyshv distance, the Manhattan distance, the Euclidean distance.

The ratio will be based on the Euclidean distance.

$$d_{ij} = \left[ \sum_{k=1}^p (x_{ik} - x_{jk})^2 \right]^{0,5} = \left[ (\mathbf{x}_i - \mathbf{x}_j)^T (\mathbf{x}_i - \mathbf{x}_j) \right]^{0,5} \quad (4)$$

We have  $n$  employees, each of them with  $k$  traits (each of them is evaluated in respect of  $k$  traits.)

Let's define:

$$x_j^* = \max_i x_{ij} \quad \text{where } j=1,2,\dots,k - \text{number of comparable traits,}$$

$$i=1,2,\dots,n - \text{number of employees.}$$

It can be then claimed that a model value is selected in every analyzed category (for each trait). Then:

$$x_1^* - \text{model value for trait no. 1}$$

$$x_2^* - \text{model value for trait no. 2}$$

$$\cdot$$

$$\cdot$$

$$x_k^* - \text{model value for trait no. k}$$

The objective of the assessment is to compare the employees in respect of several traits as the same time. By indicating the maximum value of each of the traits a certain model of employee  $x^*$  is developed with the desired model values of the traits:  $x^* = (x_1^*; x_2^*, \dots, x_k^*)$ .

Formula (4) looks as follows:

$$d_i = \left[ \sum_{j=1}^k (x_{ij} - x_j^*)^2 \right]^{0,5} \quad (5)$$

$i$  – number,  $i=1,2,\dots,n$

That value indicates the distance of the  $i$ -th employee from the comparable (model) employee, taking into account the studied and compared traits.

### Stage V

The ratio is ultimately determined as:

$$W = \frac{d_i}{\max_i d_i} \quad (6)$$

After such a transformation the ratio assumes the value from the range [0,1].

In the last stage, the values of ratio  $W$  can be used as the basis for the development of ranking of the employees.

The interpretation of the measure developed this way is as follow: the closer the value of the ratio to 1, the more different from the model (comparable) a given employee is. The best employees will be characterized by the value of the calculated ratio which is the closest to 0 – meaning that a given employee is ranked the closest to the model.

Consequently, the lower values of the suggested ratio shall mean the degree of completion of a specific objective is higher.

Furthermore, it should be noted that with the use of the information about the distance between the employee and the model in a base year, it is possible to evaluate the degree of completion of the objectives in the following years of monitoring the results.

## APPLICATION

### Study participants

The practical application of the ratio was presented with the data collected from a study on a sample of 500 employees employed as workers in a mining sector company in Poland.

### Tools

The “Your Job” questionnaire with 32 questions diagnosing 4 aspects of work was used in the study. The questionnaire is a translation of the American tool *Job Content Questionnaire – JCQ* by Robert Karasek which has been recently adapted to Poland conditions by Żołnierczyk-Zreda and Bedyńska. The questions used in the questionnaire regard: assessment of demands, decision latitude that is the feeling that the employees can meet the requirements, job insecurity and the superior’s and co-workers’ support. The responses were coded as follows: 1 - I completely disagree, 2 – I don’t agree, 3 – I agree 4 – I completely agree.

As a result of the study the traits mentioned above were measured. All calculations were made with the use of the ltm package in R program [Rizopoulos 2006].

## Results

The possibilities of use of the ratio suggested above were presented with the selected 100 employees who were diagnosed by comparing their level of the following traits: job insecurity, decision latitude and co-workers' support.

Table 1 presents only examples of estimated (with the use of R program) levels of analyzed traits in a group of 10 selected employees.

Table 1. Estimated levels of analyzed traits in a group of 10 employees

Employee	Job insecurity	Support level	Decision latitude
1	1.878	-2.830	-2.497
2	1.833	-2.382	-2.317
3	1.648	-2.056	-1.301
4	1.451	-1.501	-1.708
5	1.740	-0.389	-1.122
6	1.552	1.466	-3.042
7	1.633	-2.574	-1.430
8	1.440	-2.155	-1.393
9	1.228	-2.268	-0.835
10	1.201	-1.790	-0.175

Source: own calculation

Trait: job insecurity was changed into a stimulant by multiplying the initially estimated values by -1. Next, the maximum value was selected from among 100 employees selected for assessment for each of the comparable traits. The following values were received:

$$x_{1,\max} = 1.878 - \text{maximum value of the trait: job insecurity,}$$

$$x_{2,\max} = 2.329 - \text{maximum value of the trait: co-workers' support,}$$

$$x_{3,\max} = 1.410 - \text{maximum value of the trait: decision latitude.}$$

In the next stage, the employees were compared in respect of all three traits at the same time, calculating the value of ratio  $W$  presented for each of them. Table 2 presents the values of the ratio for the group of 10 example employees.

Table 2. Values of ratio  $W$  presented for 10 example employees

Employee	1	2	3	4	5	6	7	8	9	10
ratio $W$	1	0.928	0.797	0.765	0.574	0.702	0.876	0.819	0.796	0.689

Source: own calculation

The results obtained were the basis for the development of the ranking of 100 employees. The employees were ranked from the smallest to the biggest distance from the model (value of the ratio 1).



The conducted analysis provided the following information about the group of the assessed employees:

- in the group of 100 employees, only 10 of them demonstrated ratio  $W$  below 0.5. That means that only 10 employees met the employer's requirements/expectations at the highest degree defined as the model,
- 16 employees in the conducted assessment are the least similar to the defined model in respect of the studied traits; they met the requirements regarding the desired values of the comparable traits set by the model to the lowest degree. The value of the ratio calculated for them is higher than 0.7,
- the other employees demonstrated value of the ratio in the range [0.5; 0.7].

With the use of the ratio it is possible to select from a group of employees those who, in regard of the studied traits, hold the highest or the lowest positions. On the basis of their analysis it is possible to indicate those who change their positions over time – indicate the direction of change: positive or negative.

## CONCLUSIONS AND APPLICATION IN PRACTICE

Contemporary organizations operate in a fast-changing environment [Krupski 2005] which affects the changes in the management process. There is no universal way of managing a contemporary organization. The suggested assessment ratio facilitates the introduction of a method of monitoring the completion of specific objectives in the strategy of organization growth. The objective of monitoring is to provide a possibility of implementing actions to eliminate inadmissible deviations of the ratios from their assumed target values, to systematically monitor the completion of set objectives and to determine whether their further completion is not in danger. The suggested ratio can be used as a tool to control whether the set objectives have been achieved, to measure, analyze and evaluate work performance for specific employees. It can help organizations in achieving their long-term goals.

The potential benefits for the organizations of the use of the ratio to make assessments include the following:

- easier development of different kinds of teams of employees,
- selection of the best employees and planning their individual career paths,
- selection of the weakest employees,
- support in making decisions on how to connect the employee remunerated ratio with their work performance (pay rises, bonuses, awards),
- development of a database to facilitate assessment of progress in achieving goals set by individual employees,
- monitoring changes in case of big deviations from the required ratio values.

The assessment with the use of that kind of ratio can be used to provide the employers or managers with additional feedback information on the quality of

work of their employees and facilitate the process of identification of the factors which affect their professional growth.

## REFERENCES

- Ayala R.J. (2009) *The Theory and Practice of Item Response Theory*, Springer Street, New York.
- Jędrzejczak J. (2000) *Oceny Okresowe. Zarządzanie przez ocenianie*, Ośrodek Doradztwa I Doskonalenia Kadr, Gdańsk.
- Karasek R. (1979) Job demands, job decision latitude and mental strain: Implication for job redesign, *Administrative Science Quarterly*, 24, 285-308.
- Krupski R. (2005) *Zarządzanie przedsiębiorstwem w turbulentnym otoczeniu*, Polskie Wydawnictwo Ekonomiczne, Warszawa.
- Juchnowicz M. (2003) *Narzędzia i praktyka zarządzania zasobami ludzkimi*, Wydawnictwo Poltext, Warszawa.
- Listwa T. (1999) *Zarządzanie kadrami. Podstawy teoretyczne i ćwiczenia*, Wydawnictwo AE im. Oskara Lanego, Wrocław.
- Ludwiczynski A. (2014) *Zarządzanie Zasobami Ludzkimi. Tworzenie kapitału ludzkiego organizacji*, Wydawnictwo Naukowe PWN, Warszawa.
- Pawlak Z. (2003) *Personalna funkcja firmy. Procesy i procedury kadrowe*, Poltext, Warsaw.
- Pocztowski A. (2003) *Zarządzanie zasobami ludzkimi*, PWE, Warszawa.
- Rizopoulos D. (2006) An R Package for Latent Variable Modeling and Item Response Theory Analyses, *Journal of Statistical Software*, 175.
- Samejima F. (1997) *Graded Response Model*, Handbook of modern item response theory, Springer, New York.
- Sidor-Rządkowska M. (2000) *Kształtowanie nowoczesnych systemów ocen pracowniczych*, Oficyna Ekonomiczna, Kraków.
- Szałkowski A. (2000) *Wprowadzenie do zarządzania personelem*, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków.
- Verbruggen M. (2010) Career Counseling in the New Career Era, *Review of Business and Economics LV* (1), 2-22.
- Wilson M., de Boeck P. (2004) *Explanatory Item Response Models*, Springer-Verlag, New York.
- Żołnierczyk – Zreda D., Bedyńska S. (2014) Psychometric properties of the Polish version of Karasek's Job Content Questionnaire, *International Journal of Occupational Safety and Ergonomics*, 20 (4), 583-593.