

An Informative System Based on the Skill Gap Analysis to Planning Training Courses

Laura Antonucci¹, Francesco Domenico d'Ovidio²

¹Department of Economics, University of Foggia, Foggia, Italy

²Department of Economics and Mathematical Methods, University of Bari, Bari, Italy

Email: l.antonucci@unifg.it, francescodomenico.dovidio@uniba.it

Received October 11, 2012; revised November 7, 2012; accepted November 14, 2012

ABSTRACT

By using the skill gap analysis we propose a new quantitative method to plan training courses. By comparing the differences between the competences needed to cover a position and the ones possessed, it is possible to find the cases that need to be trained. According to different situations we propose some algorithms to plan courses and to find the subjects that require to be trained.

Keywords: Competences; Assessment; Skill Gap Analysis; Training

1. Introduction

This paper describes the basic structure of an informative system for planning courses for firms' workers, according to results of a skill gap analysis. We shall act as if we were a training company that has to planning the training of the workers of a big company, with a wide variety of competences at different levels. The system can be used to plan the training procedures, to find the persons that have to be trained and to decide the training program.

The quantitative approach proposed is quite innovative, because the literature about the skill gap analysis usually follows a qualitative approach [1,2].

After a brief description of the main definitions of competence and skills assessment (Section 2), we propose an indicator to measure the skill gap (Section 3) and some algorithms for the construction of training courses based on business requirements in different operational conditions (Section 4). In Section 5, some applications on a survey in the ICT sector (that is, Information & Communication Technology) are described.

2. Human Resources and Skill Gap Analysis

Spencer and Spencer [3] define competency as internal characteristic of an individual, that produces effective and superior performance. Spencer and Spencer's iceberg model shows relationship between competency and performance. They argue that there are surface and potential competencies. The surface competency, which consists of knowledge and skills, can be developed further by proper management. The potential competency consists of motives, traits and self-concept.

In order to improve the use of the available human capital, a competency assessment is often performed, involving the measurement of an individual's competencies and measures the cognitive ability using tests, structured interviews, job knowledge tests, diagnostic and promotion tests. The skill gap analysis could be a useful instrument to assess competences.

The gap analysis is a tool used to assess the difference (gap) between the actual state and a future goal state. So, the skill gap analysis is used to identify the skills that an individual needs, but he sometime hasn't, to carry out his or her job or to perform certain tasks effectively [4].

The skill gap analysis can be a valuable tool because it allows an employee to identify what improvements he will need to make in order to advance in his career. Perform a skill gap analysis gives to an individual an edge over the other employees who do not assess their professional development needs.

The first step in performing such analysis is to identify all the skills required by each individual to execute his/her work. It should then be possible to identify the critical and noncritical skills that are needed to carry out a role effectively. A critical skill is one that is required to complete a task successfully. A noncritical skill can enable to complete a task more quickly or efficiently, or at less cost than would otherwise [5].

There is criterion for determining whether a skill is critical or noncritical. Quite simply, the skill is noncritical if an employee lacks a given skill, but completes a task satisfactorily. Conversely, if a person completes a task but the outcome is unsatisfactory, the missing skill is critical.

By applying the skill gap analysis to the human resources of a firm, it is possible to find out which shortfalls of skill and knowledge are in an organization. It is then possible to target training resources on those necessary skills that require the most attention. This should result in the optimal use of resources in terms of improving the overall performance of the firm.

For individuals, the skill gap analysis can be used to produce personal developments and training plans, or to support appraisals.

For firms, the skill gap analysis can be used to identify which staff members have most knowledge of particular aspects of the business, as well as those with skill gaps. Furthermore, it can aid recruitment thru identifying candidates whose skills better match those needed to function effectively in a particular role. For example, in an application of such analysis to the role of a firefighter, the essential skills are: critical thinking, oral communication and the ability to work with others. Such analysis also allows the benchmarks and encourages tutoring and mentoring within teams.

The skill gap analysis can be undertaken using paper-based assessments and supporting interviews. However, if an analysis is to be performed across a large number of employees, it can create a huge burden of management and administrative. Therefore many firms use skill management software.

The analysis can be applied on a continuing basis or as a one-off exercise. Specialized software can generate a report of skill gap analysis with a few clicks of the mouse. A paper-based report takes somewhat longer, depending on the number of questions that need answers.

3. The Skill Gap Indicator

The method here proposed follows a quantitative approach and measures the gap of each competence for all the examined subjects. The procedure can be adapted to several concrete situations. As said before, we want to build a model that helps to draw a training program for the workers of a big corporation according to the needs coming from the skill gap analysis.

The algorithm will consider only the situations where a negative gap was found, that means that the competences owned by the worker are lower than the ones needed. Small and big gaps have to be distinguished: in the first case the company can organize self-training activities, but in the second case a training program is required. It is important to distinguish, also, between gaps of those competences that are often used and gaps of competences that are rarely applied.

The starting point of the algorithm is the mapping of the skill needed for each working profile, evaluated through a score or an ordinal scale. The definition of the benchmark for each competence needed can be obtained

following a top-down (from managers to employees) or a bottom up (from employees to managers) procedure.

To measure the level of competences owned by the workers considered, using the above mentioned scale again top-down or bottom up approach will be followed.

The procedure proposed has many advantages:

- 1) it does not change if top-down or bottom-up or mixed approach is used;
- 2) it gives quantitative estimations of the skill gaps, considering both importance and frequency of use;
- 3) it enables a very flexible semiautomatic procedure for planning training courses.

Thus a quantitative indicator is proposed to assess the gap between the skills required and those actually possessed by each worker, taking into account the frequency with which it is expected that those skills are to be used.

The indicator G_{ij} of the *gap* in the j -th skill is calculated through:

$$G_{ij} = (N_{ij} - P_{ij}) \times k_j$$

$$\begin{cases} k_j = 1 & \text{if } F_{ij} > F_{0j} \text{ and } N_{ij} - P_{ij} > c \\ k_j = 0 & \text{otherwise} \end{cases}$$

where N_{ij} indicates the score of *necessity* assigned to the i -th interviewee for the j -th skill, P_{ij} indicates the score attributed to the level of the skill *possessed* by the worker, F_{ij} indicates the *frequency of utilisation* of the skill¹ by the same worker; furthermore, c is a constant (chosen to cut the effect of low or negative gaps: for example, 0 or 1) and F_{0j} indicates the frequency of utilisation that, in the ideal *benchmark* of the company, represents the threshold above which the skills shortage represents a problem to be remedied (taking into account the costs of training).

The term k_j permits to taking into account only positive gaps, setting the indicator to zero if the score assigned to the level of skill possessed is greater than the score given to that necessary ($N_{ij} > P_{ij}$) and if the frequency of utilisation of such skill is less than the predetermined threshold level.

The indicator G_{ij} allows flexibility in the planning of the interventions, sizing them on the basis of company exigencies, because it is possible to vary the threshold of the intensity of use, taking into account the interests of the company, operational situations and available budget.

Moreover, in the successive algorithms of identification of individual situations, it is possible to insert one or more filters to exclude those less serious ones on which it is best to intervene with partnering/self-training: for example, to reduce the number of employees to train, it is

¹For the calculation of N_{ij} , P_{ij} and F_{ij} , it is possible to utilize the scores attributed by the worker itself and those attributed by colleagues or by supervisors. Acting in this way, G_{ij} re-enters in the canons of 360°-assessment. In this manner, the evaluation of the *gap* would take into account both the experience of people who daily performs certain roles, and the evaluations of decision makers.

possible to exclude from the programme the workers who have effective *gaps* ranging between 1/10th and 1/5th of the maximum value (depending on the characteristics and availability of the company), providing incentives for them to undertake self-training.

Only for descriptive purposes, G_{ij} can be standardised utilising, for example, the field of variation that is simply provided by $\max(N_i)$, given that P_{ij} is null when it is greater than N_{ij} . The indicator thus takes on values included between 0 and 1, eventually subject to multiplication by 100 for easier readability.

The indicator G_{ij} can be used to obtain synthetic indicators using the arithmetic mean of the individual indicators of discrepancy corresponding to each j -th skill present in the company.

To plan courses using G_{ij} , that do not consider the different importance of the competencies could arouse some perplexity among experts of adult training. It is well known that not all skills have the same importance for the company activities, nor the same importance for all individuals engaged in such activities. The indicator is able to take into account such diversity, both in the pinpointing of ideal *benchmark* skills as in the attribution of points relating to the frequency of utilisation. A shortage of a fundamental skill, in fact, produces a high scoring, while gaps in skill rarely utilised (although being important) will obtain a zero score.

4. Building Training Courses Using the G_{ij} Indicator

Using the G_{ij} indicator it is possible to obtain precise directions for the construction of training courses aimed at the professional development of workers and increase organizational efficiency.

Given a group of employees of a company or a consortium, once excluded those who do not need training, the following solutions are available:

- 1) Micro Training Modules (Par. 4.1);
- 2) Automated Training Process Assigning Students to Each Course (Par. 4.2).

4.1. Micro Training Modules

The construction of micro modules is appropriate when a conspicuous number of employees with a gap in various skills was found. In this case, the number of students and the heterogeneity of their needs suggested the creation of separate modules for each expertise in which was highlighted the presence of a gap.

Such modules will be divided into two or (at the most) three levels, thus providing entry level, intermediate and advanced courses, depending on the score of the personal skills and those needed. In fact, where the level of skills possessed is low and high skills are required, the em-

ployee may attend entry level as well as higher level courses. However, if the necessary skills are below, just a basic course² would be recommendable.

In case some modules or classes are too crowded the organizers can expect more issues of the same course or to split such modules into several classes.

Special attention should be given to the course scheduling in order to avoid overlapping modules meant for the same student. Furthermore, the courses should be articulated by taking into account the logic of the topic.

To form classes, one should be aware of the usual organizational constraints, such as number of seats available for classrooms and laboratories (for example: no more than 30 students for classroom) or a minimum number of students required to amortize the costs of the course (at least 15 per course).

Apparently, this solution is simpler and more efficient, as well as the closest to the usual training strategies, but it presents many logistic's problems. In order to be applied, it first requires that the firm has a large number of employees in difficulty, thus ensuring each micro modules a number of students that can justify the costs. Furthermore, skills that characterize the modules should be sufficiently separate. Finally, it can be hard to identify the best sequence of micro modules, if they are too numerous.

The advantage of this solution lies in the significant reduction of business costs. By carefully organizing educational activities, each worker will be removed from his/her work activity only for a short period of time.

An algorithm that can be used for the construction of training in micro classes is described by [6].

4.2. Semi-Automatic Assignment

We are now at the most important point in the perspective of this note: the identification of the different training courses starting from the actual needs given by relevant individual gap. **Figure 1** shows one of the possible sequences of operations that, starting from the "measurements" of the competence skill gap (G_{ij} indicator) should help to formulate appropriate training solutions by using a statistical procedure to identify those contact points of which will be discussed shortly and to aggregate the potential students.

Here the algorithm is briefly described (see **Figure 1**). First of all, after the calculation of the indicator G_{ij} , the organizers need to decide the minimum gap on which to intervene (remembering that minimal discrepancies can be resolved through partnering or self-training).

Then, the training company must define the number of persons to be involved in training courses. If this number

²For simplicity, the algorithm's scheme describes only the "basic" and "advanced" courses, but the procedure can be easily modified by inserting intermediate modules.

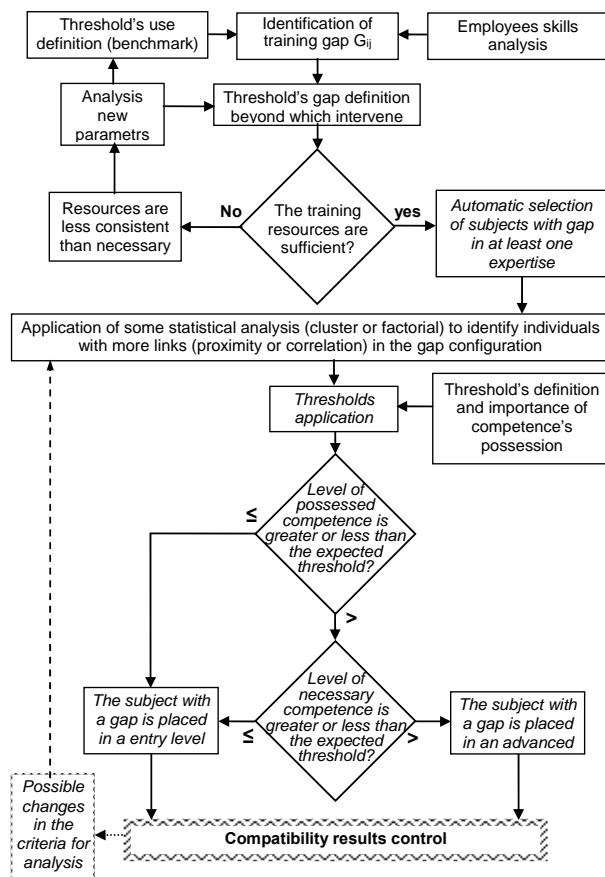


Figure 1. Algorithm for the construction of training courses by applying semi-automatic assignment of students based on statistical techniques.

is greater than the available resources, it obviously would be necessary to re-calibrate some of the threshold parameters (intensity of utilisation and *gaps*). However, if the resources are sufficient, the organizers can proceed to an appropriate statistical analysis to identify the skills most homogeneous in order to decide who has to attend the different training courses.

The statistical techniques which could be used in this case are of two types, basically: *unit classification* and *analysis of the correlation matrix* of the data.

- With regard to the first typology, a *cluster analysis* can be applied to the skills in which there is a *gap*. For example, if the skill “A” and the skill “B” both demonstrate discrepancies similar regarding the same subjects, it is appropriate to form a class in which training in both skills is provided. In the same manner, if the skill “A” reveals *gaps* in subjects 1, 2, 3 and 4, the skill “B” in subjects 3, 4 and 5, the skill “C” in subjects 2, 3 and 4, it is appropriate that the skills “A” and “C” be aggregated together.
- The second typology of statistical analysis starts with the hypothesis that, if there are groups of skills correlated to each other, the employees with significant

gaps in one of these skills may present gaps in one or more other skills related to it. To analyze the matrix of data correlations, it is possible to use the *analysis of the canonical correlations* or the *principal components analysis* and then the *factorial analysis*. In practice, applying the analysis of the principal components to all subjects with gaps, the training company can identify the most important factors in which the various required skills can be grouped. Once the articulation of subject-matter to be covered for each course is determined, the organizers can proceed to form the classroom, using the values of the skill gaps determined by the indicator G_{ij} .

After assigning subjects to different paths, there may be a need to distinguish those who have to pursue a complete course from those that can be started solely with “advanced” courses: this depends, obviously, on the level of mastery of the skills and, in the second instance, from the level of necessity for the same in the company.

It is opportune to calculate the *errors of assignment*, or quantify the training hours that employees are called on to do in matters for which no significant deficiencies emerge. It is quite clear that, the more articulated and lengthy the training courses are, the greater is the risk of dispersion of resources devoted to training. This can also be a parameter for the determination of the optimal number of training courses to be set up.

5. Method Experimentation

A procedure of automatic assigning to training courses, on the base required and actual skills, was made using the data acquired through a survey in the information technology sector in the Bari area³.

5.1. Typical Skills of the IT Sector

The survey was directed toward employees a sample of 133 IT companies located in the Bari area, with a total of 250 employees [7].

Skill gap analysis needs some information, which can be obtained either through surveys of *self-assessment*, as in the present case, or through evaluations criss-crossed between colleagues and superiors (the so-called *360°-assessment*), concerning:

- i) the company’s needs in terms of specific skills;
- ii) the store of skills specific to the subject;
- iii) the intensity and frequency of use of different skills.

The interviewees who did not provide a response to one or more requests for self-assessment of skills have been excluded from our study, as the sample group studied amounted to 221 subjects. The skills the interviewees have been asked about are those specific to the IT sector,

³For more details on this elaboration, see [7].

among which some may be considered “basic”, while others are to be considered “technical-specialist”.

The skill assessments are expressed on a scale of 1 to 5, where 1 is the lowest and 5 the highest level. For each skill, every interviewee has attributed a score both to the level of skills actually possessed and to the skill level considered necessary to carry out his/her work. Finally, a score, again on a scale of 1 to 5, has been attributed to how intensely it is utilised in the company.

The available data allow us to examine not only the deviations between the level of skills possessed and necessary by individuals, but also *how much* these skills are utilised in the company. In **Table 1**, for each skill (distinguished in “basic” and “technical specialist”), some indices of location, variability and asymmetry were summarised relating to the points attributed by the interviewees to the level of possessed skills.

In **Table 2**, statistics are shown about the level of skills considered necessary to effectuate the work⁴.

The extreme heterogeneity of the average skill level is readily deduced: in fact, the average scores vary from low values for *computer assisted design (CAD)* to values near the upper end for *browsing on the Web* or *text-processing*. For some skills of common use (such as the installation of hardware and peripherals, the use of word processors, spreadsheet, browser and multimedia programmes), the average of the scores assigned by interviewees to the level of skills possessed exceeds those needed, whereas for some more specialised skills the implementation of interventions to strengthen skills appears necessary.

Tables 1 and **2** show, also, that the index of *skewness* has frequently values very high, and the breach of the condition of symmetry could be rarely attributed to random causes⁵. The asymmetry is very strong in the case of the aforementioned expertise in *computer-assisted design*, but also in the command of *web browsing* or *word processing* skills and company needs in the *management of operating systems*. Very often, negative or positive asymmetry in the mastery of some skill corresponds to same asymmetry of its necessity of use, but not always of the same order of magnitude. In addition, for some skills the assessment presents asymmetry in one aspect and symmetry in the other. Clearly, in these cases the comparison between the mean values produces unrealistic results.

⁴The asymmetry is measured by Pearson's *skewness* index $s_k = \mu - M_3/\sigma$ which indicates perfect symmetry if it is equal to 0, while we have a positive asymmetry if $s_k < 0$ and negative if $s_k > 0$.

⁵According to the Wald test, it can be assumed that the deviation from zero is not casual (due to the probability of a α error) when the ratio $[parameter/s.e.(parameter)] > z_{\alpha/2}$, where *s.e.* indicates the standardized error found in the sample. Assuming $\alpha = 0.05$ (and so, $z_{0.025} = 1.96$) and by considering that, in the sample, the standard error of the asymmetry index is always equal to 0.164, the asymmetry should be considered not casual every time the index $s_k > 0.321$.

Table 1. Average, standard deviation and skewness index (s_k) of the scores attributed to the level of possession for some technical skills among the interviewed.

Competences/skills	Skills's possession		
	Average	Std. dev.	s_k^*
<i>Basic skills</i>			
Hardware installation	3.5	1.3	-0.47
Operating systems management	3.7	1.1	-0.60
Word processor	4.1	0.9	-0.98
Spreadsheets use	3.9	1.0	-0.71
Database use (Oracle, SQL, Access...)	3.1	1.4	-0.14
Internet navigation	4.5	0.7	-1.99
Websites creation	2.7	1.5	0.21
Multimedia (audio-video)	2.8	1.3	0.16
Installation of data network	2.9	1.5	0.04
<i>Specialistic skills</i>			
Project network infrastructure solutions	2.4	1.4	0.54
CAD/CAM/CAE aided engineering	1.5	0.9	2.33
Expertise in macro spreadsheet scripts	2.6	1.2	0.39
Data base administration (DBA)	2.8	1.4	0.07
Knowledge of Basic, Pascal, Delphi etc.	2.4	1.4	0.49
Knowledge of HTML, PHP, Java etc.	2.4	1.4	0.43
Knowledge of Assembly, C, C# etc.	2.1	1.3	0.75

*Values in italics are significant according to the Wald test, given a fixed skewness s.e. = 0.164.

Table 2. Average, standard deviation and skewness index (s_k) of the scores attributed to the level of necessity for some technical skills among the interviewed.

Competences/skills	Skills's Need		
	Average	Std. dev.	s_k^*
<i>Basic skills</i>			
Hardware installation	3.3	1.3	-0.31
Operating systems management	3.8	1.1	-0.93
Word processor	3.7	1.0	-0.47
Spreadsheets use	3.8	1.0	-0.60
Database use (Oracle, SQL, Access...)	3.8	1.2	-0.77
Internet Navigation	3.7	1.1	-0.51
Websites creation	2.9	1.4	0.12
Multimedia (audio-video)	2.6	1.3	0.37
Installation of data network	3.4	1.4	-0.45
<i>Specialistic skills</i>			
Project network infrastructure solutions	2.8	1.5	0.24
CAD/CAM/CAE aided engineering	1.6	1.1	1.79
Expertise in macro spreadsheet scripts	2.7	1.3	0.24
Data base administration (DBA)	3.4	1.4	-0.54
Knowledge of Basic, Pascal, Delphi etc.	2.6	1.5	0.25
Knowledge of HTML, PHP, Java etc.	2.9	1.4	-0.01
Knowledge of assembly, C, C# etc.	2.4	1.4	0.49

*Values in italics are significant according to the Wald test, given a fixed skewness s.e. = 0.164.

5.2. Results of the Skill Gap Analysis

Table 2 highlights the distribution of levels of discrepancy found in the sample for each of the skills considered (G_j), with fixed threshold levels for all the skills⁶. The “conditioned discrepancies” are numerically insignificant, and almost always concentrated at the low end of the distribution. There are 46 interviewees with a gap in one skill (*database use*), but none of them has gaps in a adjacent skill (*database administration*).

Table 3 does not provide information on the effective number of potential users of a training course, inasmuch as the differences could all be concentrated in the same 46 individuals (which is the maximum number of gaps found for a single skill), or be uniformly distributed in the sample, so that everyone of the 221 subjects could find themselves included in a refresher course.

The correct application of the method requires the exclusion of interviewees that show *for each skill* a number of gaps below the prefixed intervention threshold: in the considered database 112 cases are found with the indicator G_{ij} is not greater than 1 for each of the skills under study, thus the totality of subjects that should be trained amounts to 109 persons.

The selection of participants was made using *cluster analysis* applied to the skills on the basis of individual results, using the method of Ward [8] applied to the distance between the *gap* indicators, because it gave the best partition.

As shown in **Figure 2**, two main paths can be distinguished, the first of which is in turn divided into two well defined branches: that of the basic skills and that of the specialistic skills (from which the skill of *management and design of the database* remains outside).

The second group appears less defined than the first, but if the cut of the dendrogram is suitably moved left, a specific path is shown: that of the skills relating to the IT network (*Installation of the data transmission network* and *Design of the data transmission network*); in this context, however, the *Realisation of Web sites* is found to have a semi-autonomous position, and even more detached from the others, the skill of *Database use* is revealed.

Given the wide variety of skills to teach and the large number of subjects for training, it would appear to be appropriate to distinguish the students who only need “advanced” updating from those who need only “basic” updating (in other words, apply the last part of the algorithm shown in **Figure 2**). This arrangement, combined with an appropriate organisation of courses within the path, would also permit the reduction of the number of subjects set for unnecessary training.

⁶In this application, the “minimum gap” is equal to 1; in the calculation of the gap’s indicator, in addition, the threshold score for all competences (F_{0j}) is equal to 2.

Table 3. Distribution of 109 interviewees according to level indicator G_j greater than 1, with a threshold of use (F_{0j}) of 2 and intervention threshold of 1, per skill[#].

Competences/skills	G_j level >1		
	2	3	4
<i>Basic skills</i>			
Hardware installation	9	2	1
Operating systems management	17	4	
Word processor	3	-	-
Spreadsheets use	14	1	
Database use (Oracle, SQL, Access...)	23	15	8
Internet navigation	2	-	-
Websites creation	13	7	7
Multimedia (audio-video)	9	3	2
Installation of data network	18	13	5
<i>Specialistic skills</i>			
Project network infrastructure solutions	14	8	3
CAD/CAM/CAE aided engineering	4	2	2
Expertise in macro spreadsheet scripts	17	2	-
Data base administration (DBA)	-	-	-
Knowledge of Basic, Pascal, Delphi etc.	14	2	1
Knowledge of HTML, PHP, Java etc.	15	9	1
Knowledge of assembly, C, C# etc.	9	6	1

[#]The sum of cases is greater than 109 because some employees should attend more courses.

Competences:

Basic_A: Hardware Installation
 Basic_B: Operating systems management
 Basic_C: Word processor
 Basic_D: Spreadsheets use
 Basic_E: Database use (SQL server, ecc.)
 Basic_F: Internet Navigation
 Basic_G: Websites creation
 Basic_H: Multimedia (audio, video ecc.)
 Basic_I: Installation of data network
 Spec_A: Project network infrastructure solutions
 Spec_B: CAD/CAM/CAE aided engineering
 Spec_C: Spreadsheet macro scripts expertise
 Spec_D: Data base administration (DBA)
 Spec_E: Knowledge: Basic, Pascal, Delphi etc.
 Spec_F: Knowledge: HTML, PHP, XML, Java etc.
 Spec_G: Knowledge: Assembly, C, C++ etc.

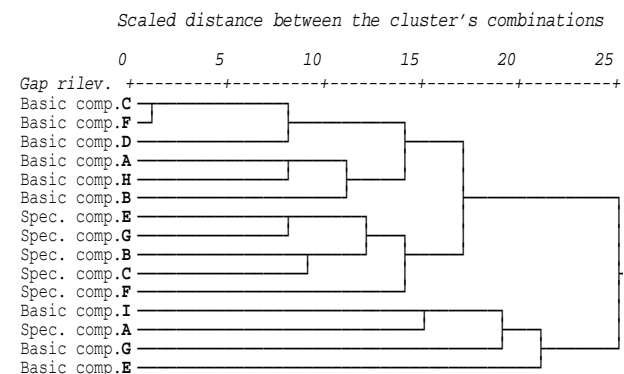


Figure 2. Cluster analysis of IT skill gaps in the Bari area.

All the paths need to be divided into several classes, since for the advanced training the numerosness in a class should be between 10 and 20 units. A number of subjects would then need to be included in more than one path: such overlaps obviously are resolved with an appropriate scheduling basis, namely, avoiding coincidence of training days. A subject might be found to follow a training course for which there is no need in two situations: the first is when he/she already possesses the knowledge and ability covered by that course; the second, when two or more courses of pertinence are interspersed with a course of no use to him/her.

Tables 4 and 5 give now a description of the results obtained by factorial analysis, with the categorical principal components method⁷ (since the evaluations are determined with ordinal scale modality, rather than that of the cardinal scale).

The Kaiser-Meier-Olkin statistic is greater than 0.7, and the selected model, as shown in **Table 4**, contains six factors, which explain almost 65% of the comprehensive variability.

The rotation by the Varimax method provides the factorial weights shown in **Table 5**, from which an appreciable separation of the identified relations is apparent (except for the *realisation of websites*).

6. Concluding Remarks

The purpose of the present work is to provide a statistical approach to the theme of the skills balance. Indeed, large manufacturing organisations, in the continuous attempt to find an organisational structure that can fully exploit available human resources, are investing heavily in the realisation of inventories and balances as to skills.

The approach proposed here is based on assessment of the gaps between the skills that a person, as a candidate to fill a particular position, should have and those that are actually possessed. Generally, this assessment has descriptive purposes only, with all the limitations arising there from: case-by-case analysis, subjectivity of assessments, lack of comparability with similar situations. The proposed solution tries to overcome the limitations outlined above with a quantitative approach, where the involvement of experts is always determinative, but the assessments of a subjective type are minimised.

The possibility to perform a gap analysis based on scores permits to comparing situations also quite different, monitoring over time the evolution of individual situations and measuring the deviations.

Given very complex situations (in terms of heterogeneity of the subjects to be trained, of the issues to be dealt with and the numerosness of the reference popula-

⁷The categorical principal components analysis, or CatPCA [9,10], is an extension of the multiple correspondence analysis, based on the AL-SOS-Alternative Least Squares Optimal Scaling.

Table 4. Number of factors relating to the skill gaps and the total variance explained ($n = 109$ subjects).

Components	Original Solution (Not Rotated)			Rotated Solution (Varimax Method)		
	Eigen values	% Variance	% Cumul. Variance	Eigen values	% Variance	% Cumul. Variance
1	2.59	17.28	17.28	2.00	13.30	13.30
2	1.87	12.45	29.73	1.94	12.89	26.19
3	1.57	10.44	40.17	1.79	11.92	38.11
4	1.32	8.78	48.95	1.41	9.37	47.48
5	1.26	8.41	57.36	1.40	9.31	56.79
6	1.10	7.36	64.72	1.19	7.93	64.72
7	0.86	5.70	70.42			
8	0.78	5.20	75.62			
9	0.70	4.67	80.28			
10	0.69	4.60	84.88			
11	0.62	4.11	88.99			
12	0.55	3.64	92.63			
13	0.45	2.97	95.60			
14	0.35	2.35	97.95			
15	0.31	2.05	100.00			

Table 5. Factorial analysis of the skill gaps in the area of Bari.

Competences/Skills	Rotated components (Varimax)					
	1	2	3	4	5	6
<i>Basic skills</i>						
Hardware installation	0.76	0.04	0.22	-0.02	0.16	0.01
Operating systems managem.	0.65	-0.10	-0.22	0.14	0.11	0.08
Word processor	0.01	0.06	0.01	0.82	-0.17	0.06
Spreadsheets use	0.00	-0.03	-0.08	0.70	0.29	-0.12
Databases (Oracle, SQL etc.)	0.09	0.03	0.06	-0.07	-0.14	0.84
Internet Navigation	0.11	0.11	-0.14	0.01	0.75	-0.17
Websites creation	0.03	0.06	0.55	0.14	0.58	0.20
Multimedia (audio-video)	0.27	-0.08	0.72	-0.16	0.12	0.01
Installation of data network	0.78	0.04	0.11	-0.05	-0.06	0.02
<i>Specialistic skills</i>						
Project network infrastr. solut.	0.49	0.25	0.13	-0.12	-0.16	-0.45
CAD/CAM/CAE aided engin.	0.03	0.25	0.54	-0.09	-0.11	0.18
Spreadsheets macro scripts	-0.10	0.17	0.64	0.14	-0.18	-0.27
Knowledge: Basic, Pascal etc.	0.04	0.76	0.19	0.10	0.17	-0.06
Knowledge: HTML, PHP etc.	-0.05	0.69	-0.01	-0.20	0.31	0.09
Knowledge: Assembly, C etc.	0.06	0.79	0.08	0.10	-0.23	-0.07

tion), use a quantitative indicator and statistical techniques to reduce the complexity of the phenomenon and to group similar cases, can be useful both in terms of reducing costs and delays in completing the analysis and in terms of the analytical nature of the results.

The reduction in the level of subjectivity of the assessments, on the one hand, represents an advantage; in the other hand, however, it exposes to some risks⁸.

In the various proposed approaches, complex techniques are utilised, usually reliable, which presuppose a good knowledge of the phenomena being studied and that the starting hypotheses are reflected in the reality. For this reason, before making a final plan of continuing education based on the proposals set out above, it is necessary to provide a series of consistency checks before and after the various stages of the process.

Ultimately, the procedure of automated designing and construction of training activity with the method of cluster analysis, proposed in these pages, has provided quite satisfactory results, while the solution relating to the use of factorial analysis had results somewhat less convincing. This may be due to the peculiarities of the totality of the skills taken into consideration and to the particularity of the sample of interviewees, but it could be due simply to characteristics of the algorithm utilised, since the cluster analysis is a procedure oriented precisely on the search for homogeneity of the ending groups.

Further experimentations could be able to provide more precise elements of judgement.

7. Acknowledgements

This paper is part of a research project financed by MIUR-PRIN 2005 "Modelli e metodi per abbinare profili formativi e bisogni di professionalità di comparti del terziario avanzato" (whose scientific coordinator was L. Fabbris, of University of Padua).

It comes from a conjoint work of the Authors, but the paragraphs 1, 2, 4 and 6 were written by L. Antonucci, while the paragraphs 3 and 5 by F. D. d'Ovidio.

REFERENCES

- [1] F. Draganidis, P. Chamopoulou and G. Mentzas, "An Ontology Based Tool for Competency Management and Learning Paths," 2006.
<http://imu.iccs.ntua.gr/Papers/C80-Ontology-based-Competency-Management.pdf>
- [2] Y.-S. Lee, J.-W. Cho, S.-J. Han and B.-U. Choi, "Dynamic Arrangement of Control in a Personalized Learning System Based on User Competency," *Communications in Computer and Information Science*, Vol. 261, 2011, pp. 468-474. [doi:10.1007/978-3-642-27180-9_57](https://doi.org/10.1007/978-3-642-27180-9_57)
- [3] L. M. Spencer and S. M. Spencer, "Competence at Work," John Wiley & Sons, New York, 1993.
- [4] H. J. Holzer, "Is There a Gap between Employer Skill Needs and the Skills of the Work Force?" In: A. Lesgold, M. J. Feuer and A. M. Black, Eds., *Transitions in Work and Learning: Implications for Assessment*, National Academy Press, Washington DC, 1997, pp. 6-33.
- [5] M. Nakayama and N. Sutcliffe, "Managing IT Skills Portfolios: Planning, Acquisition and Performance Evaluation," Idea Group Publishing, Hershey, 2005.
- [6] C. Crocetta and F. D. d'Ovidio, "Progettazione di Percorsi Formativi Mediante Skill Gap Analysis," In: L. Fabbris, Ed., *Definire Figure Professionali Tramite Testimoni Privilegiati*, CLEUP Publisher, Padova, 2008, pp. 195-214.
- [7] F. D. d'Ovidio and P. Soletti, "Adeguatezza Della Formazione Ricevuta dai Lavoratori del Settore Informatico: Analisi in Base Alle Competenze Finali," In: E. Toma and F. D. d'Ovidio, Eds., *Attività e Competenze nel Settore dell'Informatica*, CLEUP Publisher, Padova, 2008, pp. 39-68.
- [8] J. H. Ward Jr., "Hierarchical Grouping to Optimize an Objective Function," *Journal of American Statistical Applications*, Vol. 58, No. 301, 1963, pp. 236-244.
[doi:10.1080/01621459.1963.10500845](https://doi.org/10.1080/01621459.1963.10500845)
- [9] J. De Leeuw, "Nonlinear Principal Component Analysis," *COMPSTAT Proceedings in Computational Statistics*, Physica Verlag, Vienna, 1982, pp. 77-89.
- [10] J. De Leeuw and J. J. Meulman, "Principal Component Analysis and Restricted Multidimensional Scaling," In: W. Gaul and M. Schader, Eds., *Classification as a Tool of Research*, North Holland, Amsterdam, 1986, pp. 83-96.

⁸Many experts view with suspicion the use of an automated procedure for making choices that usually are made on the basis of experience (by taking into account a series of factors such as the students expectations, perspective of personal grow etc.). On the other hand, the possibility to use more objective procedure allows to "take responsibility away" from the company, but most of all to build individual path in situations that due to its complexity, would make prohibitive a careful analysis of the competences.