

VALIDATION OF A RUBRIC TO ASSESS INNOVATION COMPETENCE

Frances Watts¹, Juan A. Marin-Garcia², Amparo García Carbonell¹, Lourdes E. Aznar-Mas¹

¹ Dpto. de Lingüística Aplicada. Universitat Politècnica de València. Camino de Vera S/N 46022 Valencia. fwatts@upv.es

² ROGLE. Dpto. de Organización de Empresas. Universidad Politécnica de Valencia. Camino de Vera S/N 46021 Valencia. jamarin@omp.upv.es.

Abstract: This paper addresses the development and validation of rubrics, materials and situations for the assessment of innovation competence. Research was carried out to verify the viability of the first draft of the assessment criteria, which led to refinement of the criteria and proposals to enhance the ensuing validation process that will include students and raters of different language backgrounds.

Keywords: innovation competence; assessment; rubrics

1. Introduction

The European model of curriculum design is based on competences, although little research can be found on the key competences that a person should possess to be considered innovative. Therefore, it would be desirable to have models that enable the description of how innovation competences are developed and attained (Marín-García et al., 2011).

The Community Innovation Surveys used in the European Union (EU) derived from the guidelines to coordinate innovation research in the EU that were defined in the Oslo Manual, first published in 1993 by the Organization for Economic Cooperation and Development (OECD). The surveys highlight the importance that the concept of innovation has in the professional contexts of the majority of university degree programs. Innovation should generate competitiveness in business through the development and management of knowledge.

This paper stems from the conception of innovation as a construct (Figure 1) that is based on individual, interpersonal and networking aspects, following the model proposed by Penttilä et al. (2011, 2012).

Figure 1. Model of innovation competence construction based on Penttilä et al. (2011; 2012)



In order to acquire innovation competence it seems necessary to experiment with new teaching alternatives associated with active methodologies. However, it is also necessary to

develop new assessment methods that are valid and that produce reliable results; this calls for the definition of the learning objectives and outcomes that are to be achieved. The first step is to define the skills and capacities that make up innovation competence and, subsequently, create an instrument that will measure those skills and capacities.

With regard to the certification of the components of innovation competence, no formal system has been regulated to acquire and assess the skills and capacities involved. For example, within the higher education framework, records of the acquisition and development by students of generic competences are in want of systematic evidence. In the majority of cases, the acquisition of skills, knowledge and attitudes are inferred but not assessed appropriately, even when included as part of summative assessment.

The lack of a formal system of identification and validation of innovation competences can probably apply to other generic competences that are included in current new degree programs. For this reason, the main objective of this paper is to present a first draft of an innovation competence barometer, i.e., a rubric that comprises the cataloging and assessment of the skills and capacities that constitute innovation competence.

2. Methodology

The proposal and validation of the assessment rubric for innovation competence presented in this paper follows the steps set forth in the Instrument Development and Construct Validation (IDCV) methodology informed by Onwuegbuzie et al. (2010). The first five of the ten steps posed have been completed and are addressed here.

1. Conceptualize the construct of interest
2. Identify and describe behaviors that underlie the construct
3. Develop initial instrument
4. Pilot-test initial instrument
5. Design and field-test revised instrument

To define the construct, the items in the rubric were initially extracted from in-depth interviews with three human resource managers from different firms well known for their innovation. In the second and third steps, the items were expanded after meeting with a focus group of 12 academics. The criteria were then revised taking into account the 9 generic competences defined by the Accreditation Board for Engineering and Technology (ABET), which have been adopted in the OECD initiative for the Assessment of Higher Education Learning Outcomes (AHELO), competences that have been described by many authors (McGourty et al., 2002; Parsons et al., 2005; Passow, 2012; Shuman et al., 2005; Villa et al., 2007; Marin et al., 2011; Garzón, 2010; Penttilä et al. 2011, 2012; or Montero, in press).

In the fourth step, the first list with 39 items was sent for review to 20 academic practitioners with experience in assessment. In a subsequent plenary session with 19 raters, the analysis, annotation and filtering of the items took place. After discussion of the quality and operability of the criteria, the number of items was reduced from 39 to 25, grouped in the three dimensions shown in Figure 1. There were thus 12 items in the individual dimension, 8 items in the interpersonal dimension and 5 items in the networking dimension.

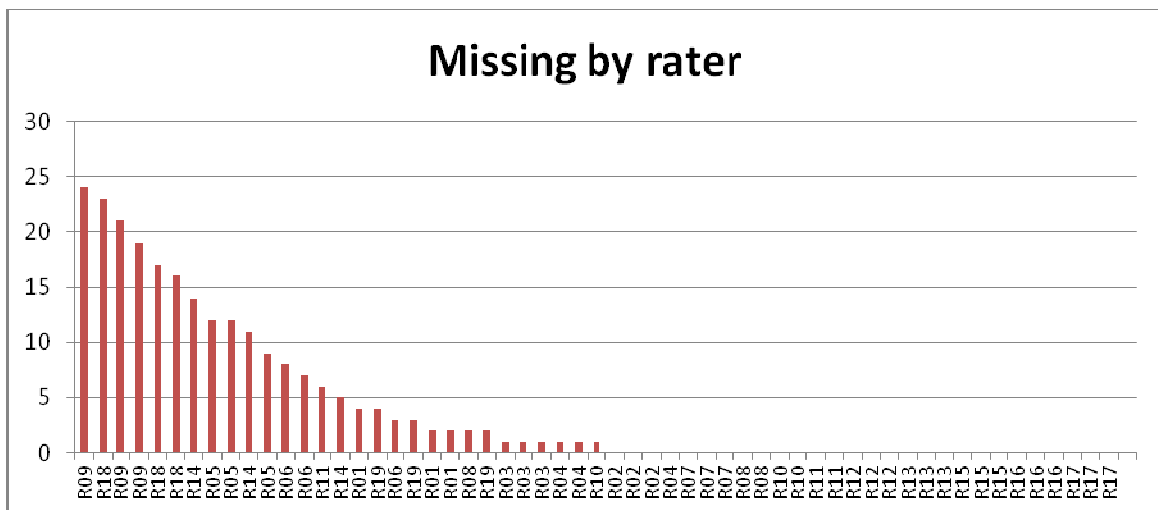
All 19 raters, of varying expertise, simultaneously participated in the trial rating of a video recording of three students who had been placed in a situation with a task that would require a display of innovation in finding a solution. Raters had 3 choices; yes, student behaviour is observed; no, student behaviour is not observed; or not applicable (n/a), student behaviour cannot be assessed because there is no evidence. In the ensuing statistical analysis, items that were left blank were counted as missing marks.

This first testing session of the instrument produced 57 ratings, i.e., 3 students were assessed by 19 raters (R01...R19). Descriptive analysis of missing marks and score frequencies led to an initial filtering of items and raters (Doval Dieguez y Viladrich Segués, 2011; Onwuegbuzie et al., 2010; Viladrich Segués y Doval Dieguez, 2011). Multiple correspondence analyses identified possible groupings of variables according to score frequency. The objective of this technique is the reduction of the dimension of a set of categorical variables, each of which consists of two or more categories (in this case, three categories: yes, no or n/a). The technique analyzes the relationship between the different categories of variables and yields as a result a bidimensional diagram. The position of the variable category in the diagram is essential, since proximity indicates relationship or association between variable categories, while distance or separation indicates the lack of a relationship or association (Greenacre, 2008; Lizasoain & Joaristi, 2012; SPSS-inc., 1990).

3. Results

In the analysis of missing ratings (Figure 2), it can be observed that raters R09, R18, R05, R14 and R06 systematically left a high number of items blank when assessing the three students.

Figure 2. Number of missing ratings classified by rater



Upon analyzing the results by item (Table 1), the items with the most missing ratings all belonged to the individual dimension.

- 05 Critically evaluates the fundamentals of ideas/actions
- 11 Uses resources ingeniously
- 12 Foresees how events will develop
- 16 Takes intelligent risks
- 17 Orients task towards target

However, regarding the items that could not be assessed in the situation that was used in the video recording (n/a), the items belonged to the networking dimension, with the exception of item 37 “Speaks foreign languages”. The items with the most n/a were:

- 32 Applies ethical values
- 33 Is able to work in cooperation in multidisciplinary/multicultural contexts
- 38 Maintains relationships with all actors engaged in a local, regional or international endeavor
- 39 Knows where to go or whom to involve to overcome difficulties and to solve problems

The data in Table 1 show that the items can be grouped in three categories (Table 2). First, there are items that cannot be observed in any of the students in the situation that was assessed. Those items should not be discarded but rather could be used in sessions for self-assessment. Or the items could be used in testing situations that favored the appearance of the behavior described in the items.

The second category contains items that are observable but do not discriminate among the individual students. This could be because the items assess behavior that is present in any student, in which case it makes no sense to include these items in the criteria. Another possibility is that the three student participants happen to be similar in these aspects. Some of these items will have to be eliminated in further analyses because they are constant.

The third category includes the rest of the items, which proved to be observable in the testing situation that was used and at the same time were discriminatory in the score totals of the students rated.

Table 1. Distribution of score frequency by item and student (A1, A2, A3)

Sum of marks in the column		275			342			162			
Student→		A1			A2			A3			
Cod	Item description	Yes	No	n/a	Yes	No	n/a	Yes	No	n/a	Missing
ind_it02	Proposes suggestions/alternatives appropriate to the demands of the task	18	0	0	17	0	0	17	2	1	2
ind_it03	Offers ideas that are original in content	17	0	1	17	0	1	9	6	2	4
ind_it04	Offers ideas that are original in the way to materialize them	14	1	3	13	1	3	6	5	6	5
ind_it05	Critically evaluates the fundamentals of ideas/actions	6	5	4	10	0	6	1	8	6	11
ind_it06	Identifies relational patterns in the components of the task	11	3	2	14	0	4	2	8	6	7
ind_it10	Approaches the task from different perspectives	13	2	3	15	1	1	4	7	5	6
ind_it11	Uses resources ingeniously	3	3	7	8	2	5	4	4	5	16
ind_it12	Foresees how events will develop	4	0	12	9	0	7	2	4	9	10
ind_it14	Shows enthusiasm	16	1	0	18	0	0	7	9	0	6
ind_it15	Is tenacious	11	1	5	16	0	2	4	4	9	5
ind_it16	Takes intelligent risks	6	6	4	10	1	4	1	6	9	10
ind_it17	Orients task towards target	15	1	0	16	1	0	11	3	0	10
int_it18	Transmits ideas coherently/effectively	15	1	1	16	0	1	6	7	2	8
int_it21	Listens to teammates	18	0	1	18	0	0	18	0	0	2
int_it23	Uses dialogue to establish constructive relationships	16	0	3	18	0	0	12	4	2	2
int_it24	Collaborates actively	15	1	2	16	0	2	11	4	1	5
int_it25	Contributes to group functioning	18	0	0	18	0	0	16	3	0	2
int_it26	Takes initiatives	15	0	2	18	0	0	2	12	2	6
int_it27	Moves others to act	10	7	1	18	0	0	3	11	3	4
int_it29	Confronts problems constructively in order to reach a consensus	7	2	7	10	0	5	3	2	10	11
net_it32	Applies ethical values	1	0	16	1	0	16	1	0	16	6
net_it33	Is able to work in cooperation in multidisciplinary/multicultural contexts	5	0	12	9	0	8	2	0	14	7
net_it37	Speaks foreign languages	18	0	0	18	0	0	17	0	1	3
net_it38	Maintains relationships with all actors engaged in a local, regional or international endeavor	2	0	16	2	0	15	2	0	16	4
net_it39	Knows where to go or whom to involve to overcome difficulties and to solve problems	1	2	15	17	0	2	1	1	16	2

Table 2. Classification of items

Category	Items
Not observable in the testing situation selected	32, 38
Observable but do not discriminate among students or only discriminate between 2.	02, 21,25, 37 03, 04,14,17,18, 23, 24
Observable and discriminate among the 3 students	05,06, 10,15,16, 26,27,29,11,12, 33, 39

The multiple correspondence analyses were carried out with the three types of possible response (yes, no, n/a), after eliminating the 5 raters with a high number of missing ratings. As items 21, 32 and 37 were constant and thus not used in the analyses, 22 valid items remained, resulting in a bidimensional representation with projections of 66 response categories (Figure 3 and Table 3).

Figure 3. Bidimensional representation of the multiple correspondence analyses (solution with 3 axes)

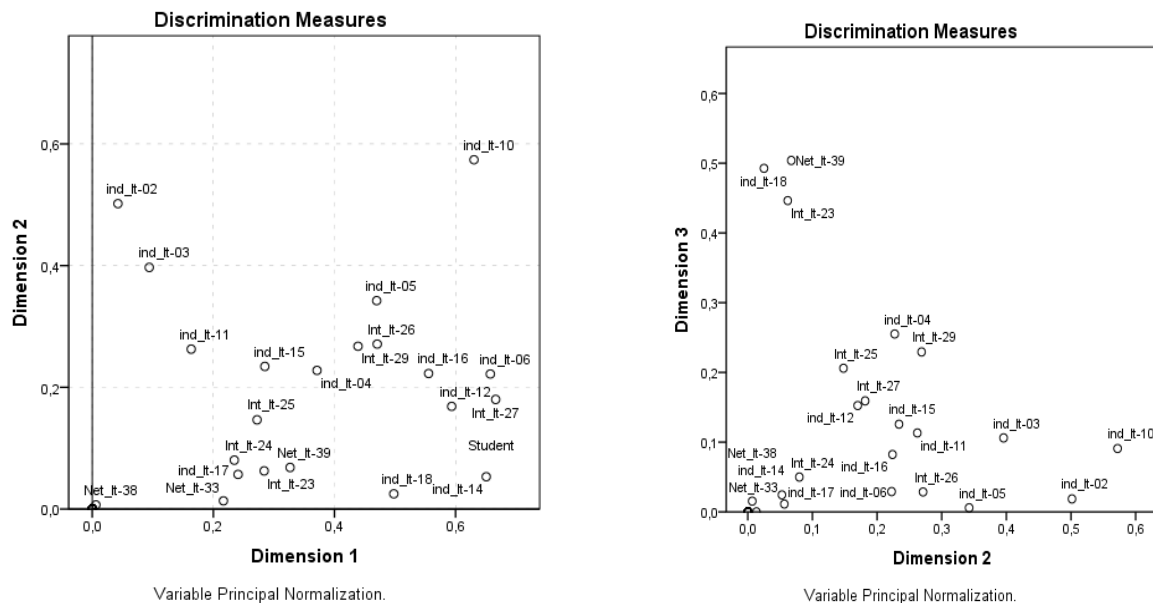


Table 3. Item discrimination

	Dimension			Mean
	1	2	3	
ind_lt02	.042	.502	.019	.188
ind_lt03	.094	.396	.106	.199
ind_lt04	.371	.227	.255	.284
ind_lt05	.470	.342	.006	.273
ind_lt06	.657	.222	.029	.303
ind_lt10	.630	.572	.091	.431
ind_lt11	.163	.262	.113	.180
ind_lt12	.594	.170	.153	.305
ind_lt14	.651	.053	.024	.243
ind_lt15	.285	.234	.126	.215
ind_lt16	.555	.224	.082	.287
ind_lt17	.241	.057	.011	.103
int_lt18	.498	.025	.493	.339
Int_lt23	.284	.062	.446	.264
Int_lt24	.235	.080	.050	.121
Int_lt25	.272	.148	.206	.209
Int_lt26	.471	.271	.029	.257
Int_lt27	.666	.181	.159	.336
Int_lt29	.439	.269	.229	.312
Net_lt33	.217	.013	.000	.077
Net_lt39	.327	.067	.504	.299
Net_lt38	.006	.007	.015	.009
Eigenvalue	8.168	4.384	3.147	5.233

The solution with three dimensions explains 71.4% of the variance in the data, with high internal consistency (Table 4). In general, the greater the distance from origin that the point representing the item is, the greater is its weight in the definition of the factor. Item 38, situated almost at the origin, is an item with little variability (almost constant) in the cases rated. The networking items have low factor loads in dimension 2 and moderate loads in dimensions 1 and 3. We suppose that with another set of data with more variability and situations more favorable to these items, the tendency would be more evident. The loads of the interpersonal items were moderate in the three dimensions, although especially so in Dimensions 1 and 3. The items termed individual have low loads in Dimension 3 and are identified as those that have high loads in Dimension 2 and almost none in Dimension 1 (items 02, 03 and 11), those that have loads to the contrary (items 12, 14 and 06) and those that are evenly distributed in both dimensions (10, 05, 04, 15, 17).

Table 4. Internal consistency of the axes in the solution with three dimensions

Dimension	Cronbach's Alpha	Variance Accounted For	
		Total (Eigenvalue)	Inertia
1	0,919	8,168	0,371
2	0,809	4,384	0,199
3	0,715	3,147	0,143
Total		15,699	0,714

4. Discussion

This paper has presented a three-category assessment model to serve as a barometer of innovation competence. An assessment rubric was created based on expert judgment and a review of the literature; the items were filtered after review by academics with experience in assessment. The paper has presented the results of the analyses carried out with the data obtained from a trial rating session with 19 raters. Results seem to indicate that there are differences between the three categories, individual, interpersonal and networking. Nevertheless, these results should be interpreted with caution since it is a pilot experience with only one recording. The need for rater training is apparent because of the missing ratings and fall-back on option n/a.

In future ratings with larger samples, items 21, 32, 37 and 38 should be monitored to see if they are constant, in order to decide on their elimination in the definitive version of the barometer. Items 12 (Foresees how events will develop) and 16 (Takes intelligent risks) seem difficult to assess, at least in short testing situation such as the one used for the trial ratings here, and may warrant elimination. Similarly, several items may be redundant. For example, item 15 (Is tenacious) seems to associate closely with item 14 (shows enthusiasm); item 18 (Transmits ideas coherently/effectively) is similar to item 23 (Uses dialogue to establish constructive relationships). Item 24 (Collaborately actively) is similar to item 25 (Contributes to group functioning). Likewise, item 37 (Speaks foreign languages) may be redundant because, in the end, for the behavior reflected in item 37 (Is able to work in cooperation in multidisciplinary/multicultural contexts), it is necessary to speak other languages to work in multicultural environments.

In conclusion, these results must be complemented with data obtained through experimentation in the classroom. For this reason, future research will focus on further field testing with students and raters of different language backgrounds in order to be able to complete the five remaining steps of the IDCV. Those steps are the following:

6. Validate revised instrument: Quantitative analysis phase
7. Validate revised instrument: Qualitative analysis phase
8. Validate revised instrument: Mixed analysis phase: Qualitative-dominant crossover analyses

9. Validate revised instrument: Mixed analysis phase: Quantitative-dominant crossover analyses
10. Evaluate the instrument development/construct evaluation process and product

Through the application of the teaching and learning methodology which has been coined “Research Hatchery” (see Genco et al., 2012, or Lehto et al., 2011), a larger, longitudinal sample will be obtained, thus allowing for the completion of the validation of the rubric.

5. Acknowledgements

This paper was written with financial support from the European Union within the Project *518132-LLP-1-2011-1-FI-ERASMUS-FEXI "INCODE- Innovation Competencies Development"*.

6. References

- Doval Dieguez, E.; Viladrich Segués, M. C. (2011). Desarrollo y adaptación de cuestionarios en el ámbito de la salud. *Laboratori d'Estadística Aplicada i de Modelització (UAB)*.
- Garzón, M. D. (2010). A comparison of personal entrepreneurial competences between entrepreneurs and CEOs in service sector. *Service Business*, Vol. 4, pp. 289-303.
- Genco, N.; Hölttä-Otto, K.; Conner Seepersad, C. (2012). An Experimental Investigation of the Innovation Capabilities of Undergraduate Engineering Students. *Journal of Engineering Education*, Vol. 101, nº. 1, pp. 60-81.
- Greenacre, M. (2008). *La práctica del análisis de correspondencias*. Fundación BBVA.
- Lehto, A.; Kairisto-Mertanene, L.; Penttilä, T. (2011). *Towards innovation pedagogy. A new approach to teaching and learning for universities of applied sciences*. Turku University of Applied Sciences.
- Lizasoain, L.; Joaristi, L. (2012). Las nuevas tecnologías y la investigación educativa. El análisis de datos de variables categoriales. *Revista Española de Pedagogía* nº. 251, pp. 111-130.
- Marin-Garcia, J. A.; Aznar-Mas, L. E.; González-Ladrón-de-Gevara, F. (2011). Innovation types and talent management for innovation. *Working Papers on Operations Management*, Vol. 2, nº. 2, pp. 25-31.
- McGourty, J.; Shuman, L.; Besterfield-Sacre, M.; Atman, C.; Miller, R.; Olds, B.; Rogers, G.; Wolfe, H. (2002). Preparing for ABET EC 2000: Research-based assessment methods and processes. *International Journal of Engineering Education*, Vol. 18, nº. 2, pp. 157-167.
- Montero-Fleta, B. (in press). Looking beyond linguistic outcomes: active learning and professional competencies in higher education. *Procedia-Social and Behavioral Sciences*. Elsevier.
- Onwuegbuzie, A. J.; Bustamante, R. M.; Nelson, J. A. (2010). Mixed Research as a Tool for Developing Quantitative Instruments. *Journal of Mixed Methods Research*, Vol. 4, nº. 1, pp. 56-78.

Parsons, C. K.; Caylor, E.; Simmons, H. S. (2005). Cooperative education work assignments: The role of organizational and individual factors in enhancing ABET competencies and co-op workplace well-being. *Journal of Engineering Education*, Vol. 94, n°. 3, pp. 309-318.

Passow, H. J. (2012). Which ABET Competencies Do Engineering Graduates Find Most Important in their Work? *Journal of Engineering Education*, Vol. 101, n°. 1, pp. 95-118.

Penttilä, T. & Kairisto-Mertanene, L. (2012). Innovation competence barometer ICB - a tool for assessing students' innovation competences as learning outcomes in higher education, in INTED2012 Conference. 5th-7th March 2012, pp. 6347-6351.

Shuman, L. J.; Besterfield-Sacre, M.; McGourty, J. (2005). The ABET "Professional skills" - Can they be taught? Can they be assessed? *Journal of Engineering Education*, Vol. 94, n°. 1, pp. 41-55.

SPSS-inc. (1990). SPSS categories. SPSS inc.

Viladrich Segués, M. C.; Doval Dieguez, E. (2011). Medición: fiabilidad y validez. *Laboratori d'Estadística Aplicada i de Modelització (UAB)*.

Villa Sánchez, A.; Poblete, M. (2007). Aprendizaje basado en competencias. Una propuesta para la evaluación de las competencias genéricas. Universidad de Deusto.