The use of fuzzy linguistic information and Fuzzy Delphi method to validate by consensus a questionnaire in a Blended-Learning environment

Jeovani Morales[1], Rosana Montes[1,2] Noe Zermeño[1], Jeronimo Duran[1], and Francisco Herrera[1,3]

[1] Andalusian Research Institute Data Science and Computacional Intelligence, DaSCI, University of Granada
{jeovani,nzermeno,jeronimoduran}@correo.ugr.es,
[2] Software Engineering department. School of Informatics and Telecommunications Engineering, University of Granada,
rosana@ugr.es,
[3] Computer Science and Artificial Intelligence department. School of Informatics and Telecommunications Engineering. University of Granada.
herrera@decsai.ugr.es,

Abstract. The virtual learning landscapes have created complex environments when evaluating an educational experience. The Fuzzy Delphi method, which incorporates the theory of fuzzy sets, takes the opinions issued by judges, from a linguistic perspective, to validate a questionnaire that will measure the degree of success of an educational experience. The judges have to reach a consensus on the validity and applicability of the instrument. This work contributes to the validation of questionnaires by enabling linguistic assessments and not only binary answers and with a calculus of consistency and consensus degrees for each item, which contributes to consensus reaching. It has been use as a practical experience to define, with the consensus of nine experts, a questionnaire that measures the virtual communication and the satisfaction with in a Blended-Learning pilot experience in the subject of Software Fundamentals, 1st semester of the Degree in Computer Engineering of the University of Granada.

Keywords: Linguistic Decision Making, Fuzzy Delphi method, B-Learning, instrument validation

1 Introduction

Newly emerging educational methodologies tend to encourage the creation of virtual learning environments. In higher education they are promoted with different technological tools for self-regulation of learning, as well as collaborative and cooperative learning [19]. Examples of this are the Flipped Classroom (FC) [18] and Mobile-Learning (ML) [11] methodologies. In addition, in the context of e-Learning, interactions between participants have been defined through the concept of Community of Inquiry (CoI) [8]. On the other hand, the application of in-person interactions combined with e-learning support is also an educational methodology that integrates the advantages offered by each of the above, and is known as Blended-Learning (B-Learning).

The application of those environments altogether become complex to evaluate as there are many constructs and different ways of structuring a measuring instrument such as the questionnaire. Moreover, specialized literature has developed each area separately, that is questionnaires for FC or questionnaires for ML or questionnaires for CoI. So there is little basis for taking a validated questionnaire to be applied in combination of the above methodologies.

Based on a proposal to evaluate a teaching experience that combines ML and FC [7] in an B-Learning environment, our aim is to know the degree of applicability of that questionnaire in a pilot educational experience, by checking the robustness of the instrument through the evaluation of judges.

In this paper, content validity of a questionnaire has been checked by the Fuzzy Delphi (FD) method, which is based on obtaining the opinion of judges in an iterative process for assessing consistency and consensus among the items of the instrument. Given that experts usually evaluate on a binary linguistic or numeric scale, the aim of this work is to use an enriched linguistic term set. In this way, we take advantage of the expert's knowledge in the assessments issued. To this end, we had the support of 9 expert judges in the area of Education Sciences and Information and Communication Technologies (ICT). At the end of the application of this methodology, a consensual questionnaire is obtained.

In the following section, a descriptive overview of the preliminaries relating to an educational experience and the validation of a questionnaire is provided. In Section 3, the FD method is applied to the research context. Finally, Section 4 presents the conclusions.

2 Preliminaries

This section describes the educational experience to be evaluated in addition to the criteria and steps required to validate data collection instrument together with the FD method.

2.1 A Blended-Learning experience in higher education

B-learning is a flexible approach to course design supported by the combination of different learning moments (face-to-face blended with online activities). Thanks to technological advances that promote interaction between students, traditional focus of education shifts from individual to collaborative approaches. Collaboration and virtual communication are fundamental aspects of e-Learning because of the effect they have on learning and satisfaction [12], so it has long been sought to analyze the characteristics necessary to improve learning outcomes in higher education environments. A theoretical and analytical model is the CoI [8] model, which is based on a collaborative-constructivist perspective of education that conceptualizes the learning and virtual communication. Thus representing the process of creating a deep and meaningful learning experience that develops through three interdependent core elements:

- *Cognitive Presence*: Through a series of phases, it allows the student to construct new educational experiences.
- *Social Presence:* Develops interpersonal relationships through the media available in the learning environment.
- *Teaching Presence:* Integrates the above elements through design, direct teaching and resource facilitation.

A relatively new and popular pedagogical methodology for B-Learning is known as*Flipped Classroom* (FC)[1]. It is based on flipping moments of learning, conceptual acquisition and application of knowledge allowing students to learn theory outside the classroom, through resources provided by the teacher, mainly videos. And also through the application of knowledge within the classroom in a collaborative and meaningful way with the support of the teacher and/or peers, promoting more active and responsible learning by students[13].

In the same sense, the use of mobile devices such as laptops or smartphones, being highly individualized and collaborative tools, has allowed the incorporation of Mobile Learning (ML) which is a methodology that intersects mobile computing with e-Learning, offering benefits for the learning environment, such as flexibility (to develop anywhere, anytime). The combined use of ML and FC methodologies enable teachers to easily provide B-Learning environments[4].

In the subject of Software Fundamentals of the 1st semester of the Degree in Computer Engineering at the University of Granada, 9 Telegram groups have been used to work with 70 students. This communication tool has made it possible to carry out synchronous meetings and asynchronous teamwork, arising from the viewing of videos and the proposal of group activities. FC has therefore been combined with ML and it is desired to evaluate the underlying CoI model in the virtual community. In order to accomplish our aim, we validate the questionnaire [7] that contains the necessary characteristics to evaluate this specificity in blended learning situations. Table 1 shows the distribution of 45 items in 7 dimensions covering the two blocks that we want to evaluate: Virtual Communication and Students' Satisfaction.

Assessment instrument in combined environments							
Blocks	Virtual Communication			Students' Satisfaction			
Dimensions	Cognitive Presence	Social Presence	Teaching Presence	Cognitive Presence	Social Presence	Teaching Presence	General Satisfaction
Items	1 - 8	9 - 14	15 - 21	22 - 28	29 - 35	36 - 41	42 - 45

Table 1. Blocks, Dimensions and Items corresponding to the questionnaire to evaluate Virtual Communication and Students' Satisfaction in FC and ML methodologies.

2.2 Instrument validation: Questionnaire

There are several methodologies for data collection, among them are the use of surveys and questionnaires. These instruments are cost-effective and time-efficient, allowing for the initiation of more developed research [14].

In order for an instrument to be valid, it must meet three requirements:

1. *Reliability*: Consisting of consistency and stability.
2. *Validity*: It is the capacity of an instrument to measure the variable for which it was designed, it contains 3 dimensions: *criteria*, *construct* y *content*. And
3. *Objectivity*: It is the degree to which this is or is not permeable to the influence of the biases and tendencies of the researcher or researchers who administer, qualify and interpret it.

Consensus is the agreement produced by consent between all members of a group or between several groups. Therefore, judgmental review process [2] is a method which reports agreement among judges regarding the evaluation of a questionnaire. According to Lynn [16] at least three judges are required for the validation of an instrument, although this is not a specific figure, it depends on the complexity of the work. In addition to, a moderator figure collects the judges' suggestions and redefine the proposal for the next iteration until a consensus is reached the the instrument can be applied. Consensus methods for questionnaire validation include the Delphi method.

2.3 Fuzzy Delphi method

The Delphi method is an iterative process [5], where participants express their opinion as many times as necessary until consensus is reached; it has the characteristic of being anonymous, thus avoiding that they are influenced by the group. The sequential process that defines the Delphi method includes three phases: (1) identify the problem and its characteristics, (2) create a coordinating group that elaborates a pilot instrument and, (3) choose the group of judges that values the instrument during iterations. Once these have been carried out, the method must go through a series of stages:

1. Disseminate the instrument to judges.
2. Sort, assess and compare the responses obtained in the first iteration.
3. Modify the instrument items according to the judges' suggestions.
4. Feedback to the judges in each iteration, at least three are recommended, or until they have a positive consistency.

At the end a report is issued describing each of the elements and stages that made up the study, the development of each iteration and the degree of consensus reached.

Ishikawa et al. [10], who introduced the FD method, argued that the classic Delphi Method requires time and high costs to achieve an efficient consensus of judges' judgments as it requires several iterations in the instrument's responses.

In addition, according to Gupta [9], in expert judgments, there is ambiguity about the different meaning or interpretation each one has of what it evaluates, so that neither real situations nor personal interpretations are usually adequately reflected by quantitative values.

Therefore the use of the FD method, which is a combination of the Delphi method and the theory of fuzzy sets proposed by Zadeh [20], solves some of the drawbacks of the classical method. It avoids confusion of common understanding between expert opinions [17] or interpretation of the responses by involving diffuse numbers and taking these opinions from a linguistic perspective, providing more reasonable results.

3 Application of Fuzzy Delphi method with linguistic assessments to get the validation of a questionnaire

A measurement instrument for B-Learning is used in Section 3.1. as exemplification for our proposal. This consist in the application of two iterations of the FD method with fuzzy linguistic information provided by our experts, as it is described in Section 3.2.

3.1 Proposal design

There are few occasions when linguistic decision-making (LDM) has been associated with the validation of a questionnaire, although it has been used to normalize the results of several questionnaires to the one given as a reference [3].

A questionnaire is defined as a set of r items $Q = \{Q_1, Q_2, \ldots, Q_r\}$ which are evaluated over q criteria $C = \{C_1, C_2, \ldots, C_q\}$ of equal weights by p judges $J = \{J_1, J_2, \ldots, J_p\}$. Judges decide if each item is valid to represent the construct for which it is designed, or should be discarded for not doing so (binary answer). To validate a questionnaire, the judges face r different decision-making problems. The assessment matrix for each item $Q_l(l = 1, \ldots, r)$, is represented by a $p \times q$ matrix. Elements are the valuation of the item over criterion C_j by the expert J_k. The full problem of questionnaire validation stores a $p \times q \times r$ matrix.

The judges answered the questionnaire using a scale of 7 linguistic terms, $S = \{s_0 = Lousy, s_1 = Very\ Wrong, s_2 = Wrong, s_3 = Moderate, s_4 = Correct, s_5 = Very\ Correct, s_6 = Excellent\}$ to express their opinion. So a single assessment over item Q_l is $s_{i_{jk}} \in S(i = 0, \ldots, 6)$. This setting completely differs from the usual binary answer (accept or discard) in the assessment matrix.

Example 1 (An Instrument to apply for B-Learning methodology).

Our aim is to use a tool that assesses the quality of virtual communication and satisfaction in higher education in combined methodologies. Thus, our problem of LDM proposes a variant of the questionnaire [7] with $r = 45$ items that are evaluated according to $q = 4$ criteria (quality, writing, belonging and scale). For the purposes of this research, $p = 9$ judges are selected considering various aspects such as: teaching experience in blended/mobile/flipped methodologies, seniority and academic degree. The semantic for the linguistic labels is $S = \{s_0 = 0, s_1 = 0.10, s_2 = 0.25, s_3 = 0.50, s_4 = 0.75, s_5 = 0.90, s_6 = 1\}$.

The instrument design is part of the FD steps as shown in Figure 1. The following sections detail the iterative processes.

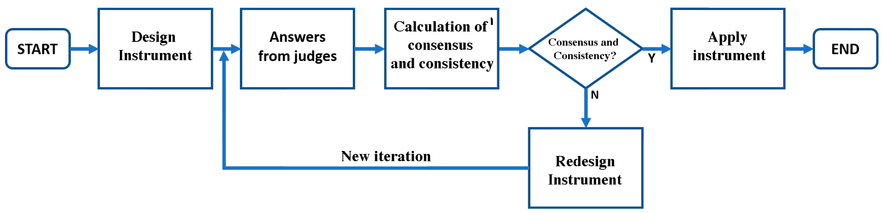

Fig. 1. Flowchart of Fuzzy Delphi methodology.

3.2 The iterative process of the Fuzzy Delphi method

This section describes the process of each iteration within the FD method used in this document. The first iteration collects the judges' opinion, calculates the consistency index and evaluates the level of consensus reached. In the second iteration, modifications to the questionnaire are made according to the judges' suggestions and disseminated together with the average response of each item so that each judge can reassess their opinion. Finally, discussion of results is made in a comparison between iteration 1 and 2.

First iteration of Fuzzy Delphi method
Once the questionnaire has been defined, it is sent to the judges for their opinion. Opinions are then represented by a family of parametric functions. Each linguistic valuation $s_{i_{jk}}$ in S is processed using a triangular function by defining a triangular fuzzy number (TFN) $s_{i_{jk}} = (a_{i_{jk}}, b_{i_{jk}}, c_{i_{jk}})$.

We consider consensus as the agreement between several members of a group. Let us note it by a boolean value CS that takes the value of T if there is consensus or F in other case. The Consistency Index $CI \in [0, 1]$ measures the degree of consensus that judges have. The closer it is to 1, the more consistent the judges' opinions are.

In our model, consistency is a boolean value noted as CC that allows us to tag the agreement as above a minimum accepted value, set as $s_4 = Correct$ within our scale. So CC take its value T when $CI \geq \varepsilon$, where $\varepsilon = 0.75$ and $\varepsilon \in [0, 1]$, since this value numerically corresponds to the s_4 label.

Example 2 (Valuations). Following our example, we represent data gather for Dimension 7 that ranges from items Q_{42} to Q_{45}. Assessments are shown in Table 2.

	Criteria															
	Clarity				**Writing**				**Belonging**				**Scale**			
	Items															
	Q_{42}	Q_{43}	Q_{44}	Q_{45}	Q_{42}	Q_{43}	Q_{44}	Q_{45}	Q_{42}	Q_{43}	Q_{44}	Q_{45}	Q_{42}	Q_{43}	Q_{44}	Q_{45}
J_1	s_4	s_6	s_6	s_6	s_4	s_6	s_6	s_6	s_6	s_6	s_6	s_4	s_6	s_6	s_6	s_6
J_2	s_3	s_5	s_4	s_4	s_3	s_5	s_4	s_3	s_3	s_4	s_5	s_4	s_3	s_5	s_3	s_6
J_3	s_5	s_4	s_4	s_5	s_5	s_4	s_4	s_5	s_4	s_4	s_4	s_4	s_4	s_4	s_4	s_5
J_4	s_6	s_6	s_6	s_6	s_6	s_6	s_6	s_6	s_6	s_6	s_6	s_4	s_6	s_6	s_6	s_6
J_5	s_2	s_4	s_3	s_2	s_2	s_4	s_3	s_3	s_3	s_4	s_3	s_2	s_2	s_2	s_2	s_2
J_6	s_3	s_4	s_4	s_4	s_3	s_4	s_4	s_4	s_3	s_4	s_4	s_4	s_4	s_4	s_4	s_6
J_7	s_3	s_4	s_4	s_4	s_3	s_4	s_4	s_4	s_3	s_4	s_4	s_4	s_4	s_4	s_4	s_5
J_8	s_4	s_6	s_4	s_5	s_4	s_6	s_4	s_5	s_4	s_4	s_4	s_4	s_4	s_4	s_4	s_5
J_9	s_6	s_6	s_6	s_6	s_6	s_6	s_6	s_6	s_6	s_6	s_6	s_5	s_6	s_6	s_6	s_6

Table 2. Judges' evaluation matrix for dimension 7 of the questionnaire (General Satisfaction).

The triangular function is applied to each valuation. Table 3 shows the valuations of item Q_{45} and its corresponding TFNs.

	Item Q_{45}: "I have a positive impression of the course."							
	Clarity		**Writing**		**Belonging**		**Scale**	
	Label	**TFN**	**Label**	**TFN**	**Label**	**TFN**	**Label**	**TFN**
J_1	s_6	(0.9, 1.0, 1.0)	s_6	(0.9, 1.0, 1.0)	s_4	(0.5, 0.75, 0.9)	s_6	(0.9, 1.0, 1.0)
J_2	s_4	(0.5, 0.75, 0.9)	s_3	(0.25, 0.5, 0.75)	s_4	(0.5, 0.75, 0.9)	s_3	(0.25, 0.5, 0.75)
J_3	s_5	(0.75, 0.9, 1.0)	s_5	(0.75, 0.9, 1.0)	s_4	(0.5, 0.75, 0.9)	s_4	(0.5, 0.75, 0.9)
J_4	s_6	(0.9, 1.0, 1.0)	s_6	(0.9, 1.0, 1.0)	s_4	(0.5, 0.75, 0.9)	s_6	(0.9, 1.0, 1.0)
J_5	s_2	(0.1, 0.25, 0.5)	s_3	(0.25, 0.5, 0.75)	s_2	(0.1, 0.25, 0.5)	s_2	(0.1, 0.25, 0.5)
J_6	s_4	(0.5, 0.75, 0.9)	s_4	(0.5, 0.75, 0.9)	s_4	(0.5, 0.75, 0.9)	s_4	(0.5, 0.75, 0.9)
J_7	s_4	(0.5, 0.75, 0.9)	s_4	(0.5, 0.75, 0.9)	s_4	(0.5, 0.75, 0.9)	s_4	(0.5, 0.75, 0.9)
J_8	s_5	(0.75, 0.9, 1.0)	s_5	(0.75, 0.9, 1.0)	s_4	(0.5, 0.75, 0.9)	s_4	(0.5, 0.75, 0.9)
J_9	s_6	(0.9, 1.0, 1.0)	s_6	(0.9, 1.0, 1.0)	s_5	(0.75, 0.9, 1.0)	s_6	(0.9, 1.0, 1.0)

Table 3. Triangular Fuzzy Numbers matrix represented as $s_{i_{jk}} = (a_{i_{jk}}, b_{i_{jk}}, c_{i_{jk}})$.

In order to find the values of consistence and consensus, we establish a conservative valuation with the triangular number $t = (l, m, u)$, set by Eq. (1). For each criterion, and with respect to the lower of the TFN experts opinions, l is the minimum value, m is the geometric mean and u the maximum value.

$$l = \min\{a_{i_{jk}}, ..., a_{p_{jk}}\} \quad , \quad m = (\prod_{k=1}^{p} a_{i_{jk}})^{(1/p)} \quad , \quad u = \max\{a_{i_{jk}}, ..., a_{p_{jk}}\} \quad (1)$$

Subsequently, the optimistic valuation $T = (L, M, U)$ is calculated using the expression given in Eq. (2) where L is the minimum value, M is the geometric mean and U the maximum value, with respect to the TFN upper values considered for each criterion.

$$L = \min\{c_{i_{jk}}, ..., c_{p_{jk}}\} \quad , \quad M = (\prod_{k=1}^{p} c_{i_{jk}})^{(1/p)} \quad , \quad U = \max\{c_{i_{jk}}, ..., c_{p_{jk}}\} \quad (2)$$

The consistency index CI is then calculated for each criterion of each item using Eq.(3). For this purpose, certain elements and requirements must be met.

In Figure 2, the elements related to the calculus of CI are appreciated and explained below.

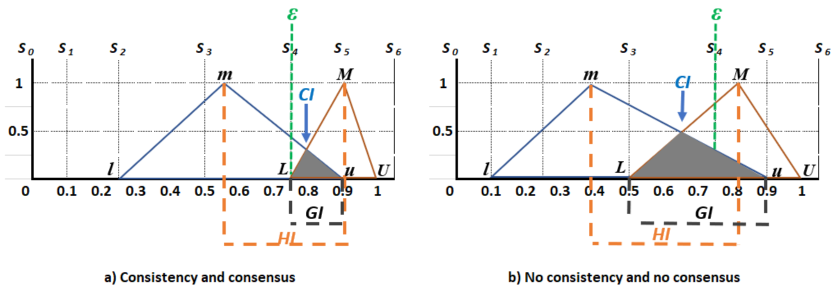

Fig. 2. Requirements for consensus and consistency. a) Representation of consensus and consistency. b) Representation of lack of consensus and consistency.

In this work we use a combination of the methods described by Dong et al [6] and Lin [15] with certain modifications, based on the following cases:

a) If $L \geq u$, the item has an excellent consensus according to the scale used, where $CS = T$,and the value of CI is:

$$CI = \frac{M + m}{2} \tag{3}$$

b) If $L \leq u$, there is a grey interval, defined as $GI = (L, u)$, one of the following 2 cases may occur:
   i) If this interval lies between the range of mean values of optimistic and conservative valuation (HI), defined as $HI = (m, M)$, consensus exists and $CS = T$, as shown in Figure 2.a). In this case, CI is determined by:

$$CI = \frac{(M \times u) - (L \times m)}{(u - m) + (M - L)} \tag{4}$$

   ii) If the GI interval is not within the HI range, there is no consistency and no consensus between the judges' valuations, as shown in Figure 2. b. Hence $CS = F$, and it is necessary to perform another iteration until all the items are consistent.

Example 3 (Consistency and Consensus). Continuing with the example of item 45, Table 4 shows the results of the application of Eq.(1) and Eq.(2).

Item Q_{45}: I have a positive impression of the course.											
Criteria	l	m	u	L	M	U	GI	HI	CI	**CC**	**CS**
Clarity	0.10	0.56	0.90	0.50	0.89	1.00	(0.50, 0.90)	(0.56, 0.89)	0.76	T	**F**
Writing	0.25	0.57	0.90	0.75	0.92	1.00	(0.75, 0.90)	(0.57, 0.92)	0.80	T	T
Belonging	0.10	0.44	0.75	0.50	0.85	1.00	(0.50, 0.75)	(0.44, 0.85)	0.63	**F**	T
Scale	0.10	0.47	0.90	0.50	0.86	1.00	(0.50, 0.90)	(0.47, 0.86)	0.68	**F**	**F**

Table 4. Content validation of Q_{45} in the first round.

The results for the criteria in item Q_{45} are as follows:

- *Clarity:* meets consistency but not consensus guidelines as the GI interval is not within the HI range, see Figure 3.a.
- *Writing*: is the only one that achieves consensus with average value of $s_4 = Correct$ and an adequate consistency, having a $CI \geq \varepsilon$, see Figure 3.b.
- *Belonging:* achieves a consensus $s_3 = Moderate$ but a consistency below the accepted value with $CI = 0.65$, see Figure 3.c.
- *Scale:* where there is no consistency or consensus, being $CI \leq \varepsilon$ and the interval GI is out of range HI, see Figure 3.d.

Concluding that the item does not have validity of content because not all criteria are satisfactorily validated. As shown in Figure 3, the consensus is between s_3 and s_4, so we proceed to make appropriate modifications based on the judges' suggestions.

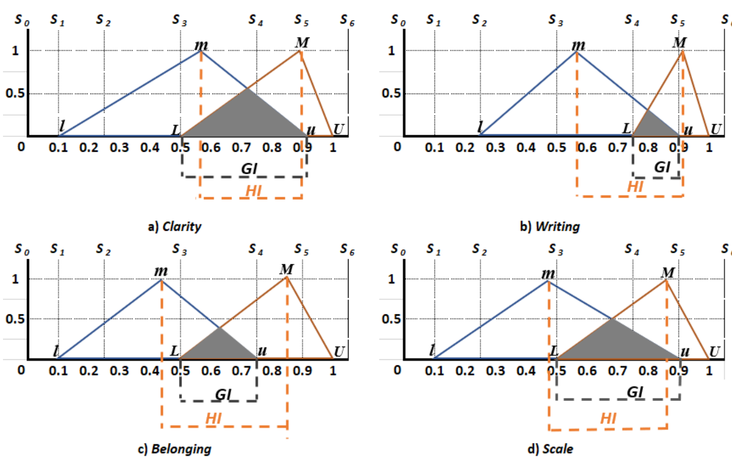

Fig. 3. The graphs represent the values for CC and CS of item Q_{45} for each criterion: a) Clarity, b) Writing, c) Belonging and d) Scale.

The moderator takes the suggestions made by the judges. In the case of item Q_{45} some of them are literally as follows: (1) *"After using the phrase* 'I am satisfied...' *in the previous questions, it change to* 'I have a positive impression...' *it breaks the dynamics of the questions of the dimension"*, (2) *"The scale I am satisfied does not agree with having a positive impression or not, it is better to have a binary scale"*. Based on these suggestions, the moderator modifies item Q_{45} *"I am satisfied with the course development"* for the second iteration.

Second iteration of Fuzzy Delphi method

Once the first iteration is completed and modifications to the first instrument have been made, the judges are provided with the average of each item obtained from the consensus reached during the first iteration. The GI and HI indices are recalculated to obtain the degree of consensus from the judges. Continuing with the example of item Q_{45}, we can see in the table 5 the results.

Item Q_{45}: I am satisfied with the course development.											
Criteria	l	m	u	L	M	U	GI	HI	CI	**CC**	**CS**
Clarity	0.90	0.90	0.90	1.00	1.00	1.00	(1.00, 0.90)	(0.90, 1.00)	0.95	T	T
Writing	0.75	0.88	0.9	1.00	1.00	1.00	(1.00, 0.90)	(0.88, 1.00)	0.94	T	T
Belonging	0.75	0.88	0.9	1.00	1.00	1.00	(1.00, 0.90)	(0.88, 1.00)	0.94	T	T
Scale	0.25	0.64	0.9	0.75	0.93	1.00	(0.75, 0.90)	(0.64, 0.93)	0.78	T	T

Table 5. Validation of the content of item Q_{45} in the second round.

The new calculations obtained from the second iteration, considerably improve what reveals that item Q_{45} has satisfactory content validity because it achieves consistency and consensus in each of its four criteria, providing an overall average for the CI of 0.91, indicating that there is a final consistency for the item of 91% agreement between the judges, thus recognizing that the FD process has been successful in this item.

3.3 Discussion of results

As can be seen in the FD process, the possibility of obtaining a consensus to validate the content of an instrument in a single iteration is complicated, whether numerically or linguistically evaluated, since each judge has their own perception of the clarity, writing, membership and scale of each item. Using linguistic terms gives the judge the ability to choose the rating that really suits his or her expertise.

The second iteration provides very significant information for judges and educational promoters of the pilot experience. It does so by highlighting the modifications, average assessments and suggestions of the judges themselves. Thus, in spite of subjectivity, it is achieved that opinions are directed towards a point in common for all, the final consensus.

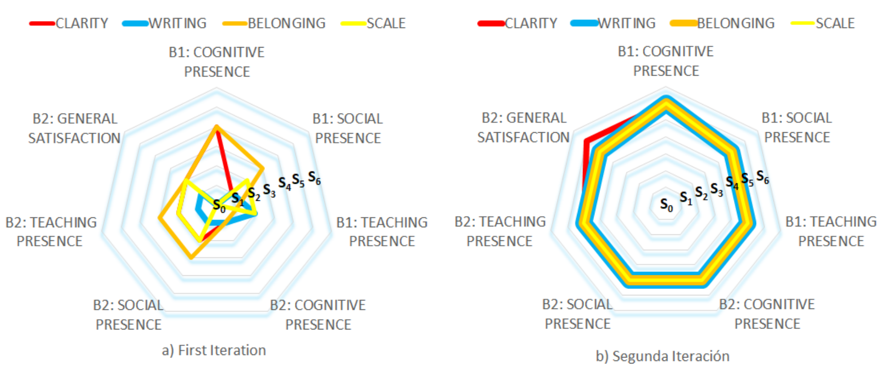

Fig. 4. Final assessment for all criteria and for all dimensions: a) First iteration. b) Second iteration.

Example 4 (Comparison of first and second iteration of FD method).

We analyze, using a radial chart, the agreement over all the dimension of the questionnaire by the application of the FD method. Results are linguistic values in S. Corresponding to the first iteration (see Figure 4.a there is no consensus neither consistency. Figure 4.b illustrates the second iteration where both criteria are satisfactorily met by having values of at least s_4.

4 Conclusions

Higher Education makes use of ICT, so it is very important to evaluate the quality of every b-Learning experience. All questionnaires should be validated prior its use. Currently, if we apply in combination the methodologies of FC and ML is hard to find in the literature a questionnaire already validated.

A questionnaire can be validated by judgmental review such as the Delphi method. This is an iterative process that tries to find consensus in the judgment opinions. When binary scales are used, much of the expert information is lost. Our proposal is to use fuzzy linguistic information to account for situations with imprecision and subjectivity. Then the use of the FD method is proposed to validate by consensus a questionnaire centered in the B-Learning environments. Our proposal incorporates the computation of the consensus status CS and consistency CC, which are fundamental to consensus reaching.

We put in practice a Fuzzy Delphi method with linguistic assessments to get the validation of a questionnaire of 45 items. The acceptance or rejection of a given item is defined as LDM problem that uses 9 judges, to assess the item considering 4 criteria. The method also uses a moderator who collects the suggestions and makes the pertinent changes for the next iteration. In our practical case, consensus was obtained with minimal loss of information due to the application of the fuzzy linguistic model.

To further validate the instrument, it will be applied with a pilot sample for statistical test, to obtaining the validity and reliability required by the instrument.

5 Acknowledgments

This document has been funded by the research project TIN2017-89517-P of the Ministry of Science and Innovation.

References

1. Bergmann, J., Sams, A.: Flip your classroom: Reach every student in every class every day. International Society for Technology in Education (2012)
2. Berk, R.: Importance of expert judgment in content-related validity evidence. Western Journal of Nursing Research **12**(5), 659–671 (1990). DOI 10.1177/019394599001200507

3. Carrasco, R.A., et al.: A linguistic multi-criteria decision making model applied to the integration of education questionnaires. International Journal of Computational Intelligence Systems **4**(5), 946–959 (2011). DOI 10.1080/18756891.2011.9727844
4. Cornelius, S., Gordon, C.: Providing a flexible, learner-centred programme: Challenges for educators. The Internet and Higher Education. **11**(1), 33–41 (2008). DOI 10.1016/j.iheduc.2007.11.003
5. Dalkey, N.: An experimental study of group opinion: The delphi method. Futures **1**(5), 408–426 (1969). DOI 10.1016/S0016-3287(69)80025-X
6. Dong, J., Huo, H.: Identification of financing barriers to energy efficiency in small and medium-sized enterprises by integrating the fuzzy delphi and fuzzy dematel approaches. Energies **10**(8), 1172 (2017). DOI 10.3390/en10081172
7. García-Lira, K., Gutiérrez-Santiuste, E., Montes-Soldado, R.: Cuestionarios para la evaluacion de la comunicacion y la satisfaccion al aplicar metodologias flipped classroom combinadas con m-learning en educacion superior. In: III Congreso Internacional de Educación Mediática y Competencia Digital, pp. 1145–1163 (2017)
8. Garrison, D., Akyol, Z.: Toward the development of a metacognition construct for communities of inquiry. The Internet and Higher Education **17**(Suppl. C), 84–89 (2013). DOI 10.1016/j.iheduc.2014.10.001
9. Gupta, R., et al.: Selection of 3pl service provider using integrated fuzzy delphi and fuzzy topsis. In: Proceedings of the world congress on engineering and computer science, vol. 2, pp. 20–22 (2010). DOI 10.1504/IJISE.2012.048862
10. Ishikawa, A., et al.: The max-min delphi method and fuzzy delphi method via fuzzy integration. Fuzzy Sets and Systems **55**(3), 241 – 253 (1993). DOI 10.1016/0165-0114(93)90251-C
11. Jaldemark, J., et al.: Editorial introduction: Collaborative learning enhanced by mobile technologies. British Journal of Educational Technology (2017). DOI 10.1111/bjet.12596
12. Kim, J.: Developing an instrument to measure social presence in distance higher education. British Journal of Educational Technology **42**(5), 763–777 (2011). DOI 10.1111/j.1467-8535.2010.01107.x
13. Lage, M.J., et al.: Inverting the classroom: A gateway to creating an inclusive learning environment. The Journal of Economic Education **31**(1), 30–43 (2000). DOI 10.1080/00220480009596759
14. Lensing, S.Y., et. al: Encouraging physicians to respond to surveys through the use of fax technology. Evaluation & the Health Professions **23**(3), 348–359 (2000). DOI 10.1177/01632780022034642
15. Lin, C.: Application of fuzzy delphi method (fdm) and fuzzy analytic hierarchy process (fahp) to criteria weights for fashion design scheme evaluation. International Journal of Clothing Science and Technology **25**(3), 171–183 (2013). DOI 10.1108/09556221311300192
16. Lynn, M.R.: Determination and quantification of content validity. Nursing research **35**(6), 382–386 (1986)
17. Noorderhaven, N.G.: Strategic decision making. Addison-Wesley (1995)
18. Tucker, B.: The flipped classroom. Education next **12**(1) (2012)
19. Valtonen, T., et al.: Perspectives on personal learning environments held by vocational students. Computers & Education **58**(2), 732 – 739 (2012). DOI 10.1016/j.compedu.2011.09.025
20. Zadeh, L.: Fuzzy sets. Information and Control **8**(3), 338 – 353 (1965). DOI 10.1016/S0019-9958(65)90241-X