Cultural adaptation and validation of the Lasater Clinical Judgment Rubric in nursing students in Spain

Montserrat Román-Cereto[a,b], Silvia García-Mayor[b], Shakira Kaknani-Uttumchandani[b], Marina García-Gámez[b], Alvaro León-Campos[b], Eloisa Fernández-Ordóñez[b], Maria Luisa Ruiz-García[b], C. Martí-García[b], Inmaculada López-Leiva[b], Kathie Lasater[c], José Miguel Morales-Asencio[b,*]

[a] Primary Health Care Centre Victoria, District of Primary Health Care of Málaga, Spain
[b] Faculty of Health Sciences, Department of Nursing, University of Malaga, Spain

[c] **Oregon Health & Science University, School of Nursing, United States**

Abstract

Background: The clinical judgment and decision-making abilities of nurses can influence many health outcomes, hence the importance of addressing these qualities in university studies. In this respect, clinical simulation is a commonly employed teaching method. The evaluation of simulation activities requires standardised instruments, such as the Lasater Clinical Judgment Rubric, which is widely used for this purpose, although a culturally adapted and validated version in Spain is not available. *Aims*: To obtain a Spanish culturally adapted and validated version of the rubric for undergraduate students of nursing. *Design, participants and setting*: Cultural adaptation and psychometric validation study carried out with undergraduate nursing students in the simulation laboratories at the University of Málaga (Spain). *Methods*: A process of translation/back-translation and cultural adaptation was carried out in accordance with international standards. The rubric was empirically evaluated in standardised scenarios with high and medium-fidelity simulators. Each student took part in two different simulation sessions, led by two instructors. In each simulation, the data were collected by two independent observers. *Results*: 152 observations were obtained from 76 students. The interobserver reliability was high, with an intraclass correlation coefficient of 0.93 (95% CI 0.92-0.95) (p=0.0001) and Cronbach's alpha of 0.93. According to the confirmatory factor analysis, the fit of the model was satisfactory in all indices, with a χ^2/df value of 1.08, GFI 0.96, TLI 0.99,

NFI 0.97 and RMSEA 0.24 (90%CI 0.000-0.066). *Conclusions*: The rubric obtained is culturally adapted to the Spanish educational context, and is valid and reliable for nursing students. Further prospective studies should be undertaken to evaluate the responsiveness, potential for transfer to clinical practice and cost-benefit ratios of different simulation designs.

Key words: Simulation Training; Patient Simulation; Education, Nursing; Students, Nursing; Clinical Competence; Educational Measurement; Psychometrics.

Introduction

Clinical judgment is the ability to make decisions based on different types of knowledge, by which the nurse can recognise significant questions, anticipate and interpret changes in a clinical situation, provide an adequate plan and implement a thoughtful intervention and reflect on its effectiveness. It also refers to the cognitive processes involved in making judgments, which includes making sense of data and cues and is defined as an interpretation of patients' health problems and needs, followed by a determined course of action (Tanner, 2006; Capelleti, Engel & Prentice, 2014). Judgment is influenced by personal experience, which contributes to the development of abilities, such as critical thinking and clinical reasoning (del Bueno, 2005). Moreover, it enables health care personnel to anticipate events and to respond appropriately to clinical situations, which may mean the difference between life and death when patients are deteriorating (or decompensating) (Alfaro-Lefevre, 2011).

Tanner described the clinical judgment of nurses as the outcome of an interpretation of the patient's needs, concerns and health problems, when the patient situation is multi-layered or unclear. Using these observations and interpretations is required to make judgments about

measures to be taken, to employ or adapt standard approaches, or to improvise those considered appropriate according to the patient's condition (Tanner, 2006), According to this model, clinical judgment should be sustained in four dimentions: effective noticing of patient problems, effective interpretation of what is noticed, an appropriate response, and effective reflection to learn from the situation; in other words, experience, skills and knowledge are essential to make appropriate clinical judgments (Martínez-Castillo & Matus-Miranda, 2015).

A systematic review was conducted to examine the findings on clinical judgment and reasoning in nursing that have emerged since Tanner's review in 2006. Electronic databases were searched to locate primary research studies about clinical judgment and reasoning in nursing. Fifteen studies were extracted and analysed, using the five main conclusions outlined by Tanner. The findings of the systematic review generally support Tanner's original model, although the importance of experience in clinical reasoning and judgment is still not well understood or fully established. In recent literature, researchers have furthered their knowledge by using tools for improving these skills in both nursing students and practicing nurses, although no strategy has been identified being more effective. This is reflected in the consideration of a sixth conclusion on clinical judgment and reasoning in nursing education strategies to improve clinical judgment, which may influence what a nurse brings to the situation (Cappelletti, Engel, & Prentice, 2014). As they become more complex (with higher patient–nurse ratios, limited clinical facilities, and shortage of nurse educators), the demands for higher cognitive skills from nurses are required (van Graan, Williams & Koen, 2016). Good clinical judgments by nurses underlie many successful health outcomes; in contrast, their absence increases the likelihood of adverse events (Aiken et al., 2014; Faisy et al., 2016; Thompson et al., 2013), Hence, it is critical to begin the development of these qualities at any early stage, preferably during academic studies. Educational interventions have been

developed with the aim of enhancing clinical judgments, but to date, the evidence is inconclusive as to which methods are best suited for this purpose (Thompson & Stapley, 2011).

In Europe, the convergence to the European Higher Education after the Bologna Process has led to a large-scale attempt to standardize education across countries, supported by pillars that focus on quality, mobility and skills oriented (Collins & Hewer, 2014). The Bologna process was a reform of higher education systems carried out in 1999 in EU countries. In the case of Spain, since 2010, Nursing Studies has become a four-year degree of 240 European credits (ECTS). This curriculum guidance is designed to train professionals to meet the challenges described above. Within this curriculum, clinical training has an important relevance, with a total of 84 ECTS through clinical placements where students are evaluated in the acquisition of their clinical competencies. This is supported by training with simulation scenarios, based on the competencies that nursing students should acquire in real clinical environments. Thus, clinical competencies get a strategic value in the development and design of Nursing Studies and they must address knowledge, skills, attitudes and values to facilitate the ability to handle ambiguous situations, tolerate uncertainty and the decision-making process with limited information (Epstein & Hundert, 2002).

Simulation is a teaching methodology in which purposefully constructed scenarios are employed to provide the representation of a real event, to facilitate practice, learning, evaluation, testing and knowledge acquisition regarding human systems, actions, or patient care. Among other purposes, it has been used in nursing education to teach psychomotor skills, to evaluate nursing competencies, and to develop clinical judgment, all in a safe environment. Recent research from a very large U. S. study has demonstrated that simulation can safely replace up to 50% of clinical education without diminishing learning or competence (Hayden et al., 2014).

The use of simulation in nurses' education is particularly valuable when psychomotor skills are to be evaluated in simulated clinical settings, with senior students and with high-fidelity simulators (Kim et al., 2016; Shin et al. 2015). Furthermore, learning through the simulation of standardised patients has achieved positive results in self-efficacy and motivation for learning (Oh, et al., 2015). However, yet. there is no solid evidence of the effectiveness of simulation in enabling nursing students to recognise when a patient's condition is deteriorating and how to respond appropriately (Fisher & King, 2013).

Background/Literature

Despite the widespread use of simulation in nursing education and notwithstanding the positive effects described above (Cant & Cooper, 2017; Hope et al., 2011; Norman, 2012), a wide range of measurement tools has been proposed in this area, many of them deficient in validity and/or reliability. Moreover, the methods used to evaluate the implementation of these instruments differ greatly, and many studies in this field present major weaknesses, such as lower level designs, or poor reporting of effect size (Cant & Cooper, 2017).

The evaluation of simulation can address many aspects, including technical skills, theoretical knowledge, confidence, attitude, feedback, self-efficacy and satisfaction (Seagull & Rooney, 2014). The use of unified and standardised instruments makes it possible to compare the performance of simulations in different situations (with different simulators, designs and instructors).

One instrument that has been widely used to evaluate the clinical judgment of students during simulation is the Lasater Clinical Judgment Rubric (LCJR) (Lasater, 2007). Based on Tanner's model of clinical judgment (Tanner, 2006), it uses Tanner's four aspects: recognition of the situation, interpretation, effective response, and reflection. It offers 11 dimensions that further describe these aspects and four levels of skill that describe each dimension, thereby forming

a trajectory of clinical judgment acquisition. Points can be assigned to each level or as a tool for offering feedback. It can be used as an observation tool by evaluators (Johnson, et al., 2014; Lasater, 2011) or completed by students as a self-assessment exercise (Cato, Lasater, & Peeples, 2009).

Several psychometric validations of the original LCJR are available (Ashcraft et al., 2013; Kim et al., 2016; Shin et al., 2014; Victor-Chmil & Larew, 2013). The Korean-language version (Shin et al., 2015) corroborated its internal consistency (Cronbach's alpha 0.86 to 0.97). However, the methods used to assess validity and reliability – especially about construct validity – have not always been the same, and robust methods such as confirmatory factor analysis have been employed in very few cases (Adamson et al., 2012; Shin et al., 2015; Victor-Chmil & Larew, 2013). No such validation study of this instrument has been carried out in Spain.

A review of simulation evaluation tools (Adamson et al., 2013) concluded that the use of existing tools, rather than developing new ones, strengthens the motivation to build on previous knowledge and highlights the importance of making greater use and further development of these instruments. Considering that clinical judgments are influenced by the context of the organization in which they take place (specific contextual factors: socioeconomic situation, political factors), the cultural adaptation of this rubric becomes necessary (Cappelleti, Engel and Prentice, 2014). Moreover, these authors stressed the benefits from validating instruments for use in contexts other than those for which they were originally designed (Adamson et al., 2013; Cappelleti, Engel & Prentice, 2014).

The aim of this study was to obtain a Spanish culturally adapted and validated version of the LCJR in university-level nursing education and to assess its implementation in the Spanish educational context.

Methods

The original LCJR was subjected to a cultural adaptation and psychometric validation study, with the participation of third and fourth-year university nursing students. First and second-year students were excluded as they do not yet have sufficient theoretical clinical knowledge with which to address the simulation scenarios presented. The third-year students took part in the pilot phases, while the fourth-year students were considered in the empirical validation.

The study participants were all fourth-year undergraduates at the University of Málaga (UMA). International exchange students were excluded because they were not Spanish native speakers. No sampling was performed, and thus the entire eligible population was included in the study.

For the translation and back-translation processes, the recommendations of the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) and the Patient-Reported Outcomes Measurement Information System (PROMIS) were followed (preparation, forward translation, reconciliation, back translation, back translation review, harmonization, cognitive debriefing, review of cognitive debriefing results and finalization, proofreading; and final report) (Wild et al., 2005) Thus, two members of the research team with good English-language proficiency (level C of the Common European Framework of Reference for Languages), working independently, performed the translation and back-translation of the original instrument. A third independent (professional) translator, a native English speaker, reconciled the two translations to create a hybrid version, stating the reasons for the modifications applied. To check that the translated version was comprehensible, a group of experts were subjected to cognitive interviews to review and identify items that might be problematic and to reach a consensus on the modifications needed. This group of experts was composed of eight lecturers and professors at the Faculty of Health Sciences, with extensive teaching, clinical, research and simulation training experience. The final version of the text was piloted in a simulation session forming part of the teaching subject “Mental Health Nursing”, by two external

evaluators, with junior students to evaluate its practicality and comprehensibility. In the cultural adaptation phase, the final translation into Spanish was reviewed by the group of experts, which consisted of eight nurses (two men and six women), all with extensive clinical, research and teaching experience, from the Nursing Department of Málaga University. The external evaluators were trained to use LCJR in a session that include comprehensive reading, brainstorming and debriefing. The original author reviewed the final version.

Data collection

The students who agreed to participate in the study (n=76) were divided into subgroups of four, who only interacted with each other during the post-simulation debriefing. Data collection took place during the sessions simulating standardised scenarios as part of the Observed Structured Clinical Evaluations (OSCE), held for fourth-year nursing degree students at the UMA Faculty of Health Sciences in April, May, and June of 2016. Each student participated in two simulation sessions with different designs: one session in an advanced life support setting and the other concerning the care of a patient with acute respiratory failure. These sessions were two months apart.

The simulations were conducted by two evaluators responsible of describing the environment and contextualising the clinical case. The evaluators were Nursing Degree Professors.

Also, two independent observers performed students' clinical judgment evaluation using the adapted LCJR. They were not able to interact directly with students, as they made their observations form a separated room, that communicated to the simulation room by one-way mirrors. This made it possible to score the results of each evaluation without any interaction with students. Once the case had been proposed, the student was allowed ten minutes to provide a response (and was given the possibility of interacting with the evaluating teachers, who then gave the standardised answers established for this purpose). Subsequently, a ten-

minute debriefing was conducted with each subgroup of four students, prompted by the two evaluators (Figure 1).

Ethical issues and permissions

This study complies with the criteria of risk-free research (it consists of the filling completion of a questionnaire), and was accepted, approved, and funded by University of Málaga (Spain), as part of the 2015-2017 Projects for Educational Innovation (PIE 144-2015).

Permission was requested from the original author to perform the psychometric validation and cultural adaptation.

The participating students did so on a voluntary basis, prior informed consent request, and all data were processed in an aggregated form and anonymised.

Analysis

Using exploratory analysis, descriptive statistics of the variables were obtained, including measures of central tendency and dispersion or percentages, depending on the nature of the variables. In all cases, the normality of the distribution was evaluated by the Kolmogorov-Smirnov test. The asymmetry, kurtosis and histograms of the distributions were also determined.

Bivariate analysis was performed by Student's t test for independent groups and for paired groups, to analyse differences between evaluations, evaluators, and gender. The chi square test was also performed. For non-normal distributions, the Wilcoxon and Mann-Whitney U non-parametric tests were conducted. In addition, bivariate correlation analyses were performed to obtain the Pearson and Spearman correlations, and the intraclass correlation coefficients were calculated.

For the psychometric evaluation, the ceiling-floor effect was analysed from the endorsement frequency of the items, taking a maximum frequency threshold of 80%, for both the lower and the upper ranges (Terwee et al., 2007). The sensitivity to change of the partial and total scores (responsiveness) was analysed by calculating the mean difference in the LCJR between the advanced life support simulation carried out in April and May and the OSCE simulation carried out in late June.

Internal consistency was calculated by Cronbach's alpha. For the construct validity, a confirmatory factor analysis was carried out, taking as indices of absolute fit of the root mean square error of approximation (RMSEA) and the goodness of fit index (GFI). For the RMSEA, 0.08 was taken as the maximum threshold of good fit. For the GFI, a corresponding value of 0.95 was established. For the incremental fit, we used the normed fit index (NFI), the Tucker-Lewis index (TLI) and the comparative fit index (CFI), with a threshold of 0.95 in every case. Finally, the parsimony of the instrument was evaluated by chi-square divided by the degrees of freedom (χ^2/df), with a threshold value of 3, and by the adjusted global fit index (AGFI), with a threshold of 0.95. Multinormality was verified by the analysis of kurtosis and the Mardia's coefficient. All analyses were performed with the statistical packages SPSS 22 and AMOS 21 (IBM Corp., 2013).

Results

The final translation into Spanish was reviewed by the group of experts: eight nurses whose mean age was 37.1 years (range 28-57), with professional experience ranging from 10 to 27 years in nursing. In this phase, only four items of the LCJR raised doubts as to their interpretability in the Spanish context and had to be discussed by the panel to agree upon a semantic acceptable interpretation.

When this semantical-adapted version was available, it was empirically evaluated on a final sample of 76 nursing students, with two observations for each of the OSCE scenarios, with independent evaluators for each simulation (Number of observations = 152). Women comprised 70.9% of the students, and 29.1% were men, with an overall mean age of 23.91 years (SD: 3.77).

The scores obtained by the students in the two evaluations are listed in Table 1. The items referring to skill and evaluation/self-analysis obtained the highest scores. No ceiling or floor effect was observed for any of the items. The “Clear communication” dimension obtained the lowest scores, which were well above the frequency of the floor effect (24.1% and 23.4% in each simulation). The values obtained for the dimension “Focused observation” were well below those of the ceiling effect (6.9% and 10.9% in each simulation).

To analyse responsiveness, we examined whether the scores varied between the evaluations made in each OSCE. There were no significant differences, except in the dimension “Making sense of data”, which were slightly higher in the second OSCE (Table 2). In addition, we tested for differences in the scores by gender; none were found (Table 3).

The correlations between the students’ ages, the values of the different items in the LCJR, and the overall scores were analysed, revealing a very modest but significant correlation between age and data prioritisation (rho = -0.168; p=0.042), and skills (rho = -0.170; p=0.039). In both cases, the correlation was inverse, i.e., the younger students obtained higher scores in these dimensions. The inter-rater reliability, evaluated by the intraclass correlation coefficient, was 0.96 (95% CI: 0.94 - 0.97) (p=0.001) for the total scale score.

The internal consistency analysis yielded a Cronbach’s alpha of 0.93. The average inter-item correlation was 0.58 (range: 0.32 - 0.85) (Table 4). The values for the homogeneity index of the scale were appropriate, above 0.5 in every case (Table 5).

Confirmatory factorial analysis (CFA) was executed, creating a factorial model that reproduced the conceptual structure on which the instrument had been originally developed, to empirically test the null hypothesis of its validity (Figure 2). The fit of the model was satisfactory in all the indices used, with a χ^2/df value of 1.08, GFI 0.96, TLI 0.99, NFI 0.97 and RMSEA 0.024 (90%CI 0.000 - 0.066).

A power analysis was carried out to test if the obtained goodness of fit was accurate. According to MacCallum et al. (1996), to test the null hypothesis R<=R0, where R is the population RMSEA (0.05) and R0 the observed RMSEA (0.024), with a type I error rate of 0.05, and taking both observed and unobserved variables for the estimation of degrees of freedom, a power of 0.81 was obtained, which is a threshold accepted as sufficient.

Discussion

The overall aim of this study was to develop a cultural adaptation and validation of the LCJR for its use in university nursing studies in Spanish, and to analyse the reliability and validity of the construct obtained, considering its extensive use in other countries.

The translation/back-translation process offered few difficulties and only four items in the translated LCJR generated discrepancies. Comparison of the empirical results obtained by the Spanish version with those of other researchers using the original LCJR suggests that the cultural adaptation did not significantly modify the original constructs. The solid theoretical basis underpinning the rubric is a good indication of its construct validity. Our results corroborate the existence of a latent factorial structure that represents these constructs in the four dimensions identified, with good indices of fit (Herrero, 2010) and high internal consistency and inter-rater reliability. A systematic review published in 2012 summarised results of the validation of LCJR in English, demonstrating a high internal validity, with a Cronbach alpha value between 0,97 and 0,81, in the included studies (Victor-Chmil & Larew, 2012). More

recently, other authors have made a validation of the original version of LCJR, obtaining Cronbach alpha values of 0.97 (Adamson et al 2012) and 0.87 (Kim et al 2016). In the Korean version, this value was 0.86 (Shin et al 2015).

Nursing education and its regulation continues to be of crucial importance because there is still too much variance of university training for nurses (National Academies of Sciences, Engineering, and Medicine, 2016), despite WHO efforts to standardise minimum regulatory requirements in this field (World Health Organization, 2009). Further research into training and evaluation methods that offer high external validity which can be used in diverse educational contexts is essential if we are to advance faster in the standardisation required of nursing education, on a global scale.

The sensitivity analysis to change in the LCJR revealed no significant differences between the scores obtained in the first and second simulation experiences, except in the dimension “Making sense of data”. This outcome was probably influenced by the short period of time between the two simulations. In the future, it would be advisable to carry out prospective studies with longer follow-up intervals to obtain more conclusive results on responsiveness.

The dimension that produced the lowest scores in the evaluation of clinical judgment was “Reflection”, which includes two dimensions: Evaluation/Self-analysis and Commitment to improvement. Evaluation of the domain “Interpretation” includes prioritising and making sense of data. In this domain, more than in others, the evaluation depends on the teacher’s interpretation of the student’s hidden knowledge (Beckham, 2013; Shinnick and Woo, 2013) it depends on what the evaluator understands that the student is interpreting, relating and analysing, i.e., the internal mental processes performed, that can be supported by a questioning process during the debriefing process intended to the student’s reflection and making sense of data. The LCJR has been used in two ways: self-administered and by the teacher. The fundamental difference between the use of the rubric in each form is probably

found in this dimension, since the student will always be aware of what s/he knows, associates and interprets, even if s/he is not able to demonstrate this objectively, facilitating evaluation. It also depends on the purpose of simulation, whether for learning (formative evaluation) or grading/qualifying (summative or high stakes evaluation).

Future research could be aimed at exploring differences in validity and reliability, whether the LCJR is administered by the teacher or is self-administered. In addition, self-administration of the LCJR could be useful in developing self-criticism, reflection and feedback on clinical practice, thus providing a useful guide for the students themselves for making mid-course corrections and setting goals (Lasater, 2011). Ashcraft et al. (2013) mentioned the possibility of using the rubric exclusively for final-year students, since the evaluation of clinical judgment is more complicated in the case of beginners.

In addition to the above suggestions, it would be useful to evaluate the validity of the rubric with different degrees of exposure to the simulation to determine its possible influence on competencies and on the constructs of the instrument. Also, it should include students' use of the rubric, as well as students' and professors' satisfaction with it.

Furthermore, the type of scenario simulated may determine the range of clinical judgment skills considered (Adamson et al., 2012). In our case, two different scenarios were presented with similar complexity, to offer a broader range of possibilities for the evaluation. The results obtained showed that these scenarios offered sufficient and adequate complexity to avoid the presence of ceiling and floor effects. Nevertheless, our study did not intend to evaluate the possible differences in clinical judgments, depending on the use of high or low-level simulation methods. This issue should be tested in further studies.

We controlled for the possible influence of the competence of the evaluators by selecting evaluators with ample teaching experience, and thus acceptable inter-rater reliability was

obtained. Test-retest was not evaluated in this study, due to the time interval between measurements that could have influenced the students' scores because of their learning experiences.

Finally, in this study it was not possible to analyse the impact on transference to competence in real practice, i.e., whether the results obtained by the students in a simulated test can be considered a valid, reliable measure of their competence in a clinical setting. Some studies have attempted this (Hayden et al., 2014; Lasater, Johnson, Ravert & Rink, 2014), this issue. For example, Hayden et a. evaluated the impact of substituting high-quality simulation for up to half of traditional clinical placements in students, with no difference on clinical competence and readiness for practice during their first six months as graduated nurses. Nevertheless, this issue should be explored further in future research, in different settings and samples, this question should be explored further in future research. It is also important to note that no rubric by itself can replace the set of factors involved in clinical judgment, and therefore, instruments such as the LCJR should always be used as a complementary tool, together with other evaluation mechanisms (Lasater, 2011).

Conclusions

It is essential to use valid, reliable methods to evaluate clinical learning, and simulation can play a valuable role in this area. This study describes a version of the LCJR adapted to the Spanish educational context, reproducing the conceptual constructs underlying the original design, and achieving appropriate validity and reliability. In addition, we present concepts and terms facilitating the standardisation of such evaluation processes and enabling feedback to students.

Nevertheless, further prospective studies are needed to evaluate the impact of simulated learning with respect to contexts of actual clinical practice and to determine the evaluation of

responsiveness over longer periods of time. Finally, it would be of interest to carry out a cost-efficiency analysis, since learning by simulation requires an investment of resources, which must be justified by rigorous scientific study.

REFERENCES

- Adamson, K.A., Gubrud, P., Sideras, S., Lasater, K., 2012. Assessing the reliability, validity, and use of the Lasater Clinical Judgment Rubric: three approaches. J. Nurs. Educ. 51, 66–73. doi:10.3928/01484834-20111130-03
- Adamson, K.A., Kardong-Edgren, S., Willhaus, J., 2013. An updated review of published simulation evaluation instruments. Clin. Simul. Nurs. 9, e393–e400. doi:10.1016/j.ecns.2012.09.004
- Aiken, L.H., Sloane, D.M., Bruyneel, L., Van den Heede, K., Griffiths, P., Busse, R., Diomidous, M., Kinnunen, J., Kózka, M., Lesaffre, E., McHugh, M.D., Moreno-Casbas, M.T., Rafferty, A.M., Schwendimann, R., Scott, P.A., Tishelman, C., van Achterberg, T., Sermeus, W., RN4CAST consortium, 2014. Nurse staffing and education and hospital mortality in nine European countries: a retrospective observational study. Lancet 383, 1824–1830. doi:10.1016/S0140-6736(13)62631-8
- Alfaro-Lefevre, R., 2011. Critical Thinking, Clinical Reasoning, and Clinical Judgment: A Practical Approach, 5th edition. Saunders, St. Louis, MO.
- Ashcraft, A.S., Opton, L., Bridges, R.A., Caballero, S., Veesart, A., Weaver, C., 2013. Simulation evaluation using a modified Lasater Clinical Judgment Rubric. Nurs. Educ. Perspect. 34, 122–126.

— Beckham, N.D., 2013. Objective structured clinical evaluation effectiveness in clinical evaluation for family nurse practitioner students. Clin. Simul. Nurs. 9, e453–e459. doi:10.1016/j.ecns.2013.04.009

— Cappelletti A., Engel J., Prentice D.2014. Systematic Review of Clinical Judgment and Reasoning in Nursing. J Nurs Educ. 53(8) 453-458. doi: 10.3928/01484834-20140724-01)

— Cant, R.P., Cooper, S.J., 2017. Use of simulation-based learning in undergraduate nurse education: An umbrella systematic review. Nurse Educ. Today 49, 63–71. doi:10.1016/j.nedt.2016.11.015

— Cato, M., Lasater, K., & Peeples, A. I. (2009). Student nurses' self-assessment of their simulation experiences. Nursing Education Perspectives, 30(2), 105-108.

— del Bueno, D., 2005. A crisis in critical thinking. Nurs. Educ. Perspect. 26, 278–282.

— Collins, S, Hewer, I. 2014. The impact of the Bologna process on higher nursing education in Europe: A review. International Journal of Nursing Studies 51, 1, 150-6

— Epstein, R. M., Hundert, E. M. 2002. Defining and assessing professional competence. JAMA: The Journal of the American Medical Association, 287, 2, 226–235

— Faisy, C., Davagnar, C., Ladiray, D., Djadi-Prat, J., Esvan, M., Lenain, E., Durieux, P., Leforestier, J.-F., Marlet, C., Seijo, M., Guillou, A., 2016. Nurse

workload and inexperienced medical staff members are associated with seasonal peaks in severe adverse events in the adult medical intensive care unit: A seven-year prospective study. Int. J. Nurs. Stud. 62, 60–70. doi:10.1016/j.ijnurstu.2016.07.013

— Fisher, D., King, L., 2013. An integrative literature review on preparing nursing students through simulation to recognize and respond to the deteriorating patient. J. Adv. Nurs. 69, 2375–2388. doi:10.1111/jan.12174

— Hayden, J. K., Smiley, R. A., Alexander, M., Kardong-Edgren, S., Jeffries, P. R. (2014). The NCSBN National Simulation Study: a longitudinal, randomized, controlled study replacing clinical hours with simulation in prelicensure nursing education. Journal of Nursing Regulation, 5(2), C1-S64.

— Herrero, J., 2010. El análisis factorial confirmatorio en el estudio de la estructura y estabilidad de los instrumentos de evaluación: Un ejemplo con el Cuestionario de Autoestima CA-14. Psychosoc. Interv. 19, 289–300.

— Hope, A., Garside, J., Prescott, S., 2011. Rethinking theory and practice: pre-registration student nurses experiences of simulation teaching and learning in the acquisition of clinical skills in preparation for practice. Nurse Educ. Today 31, 711–715. doi:10.1016/j.nedt.2010.12.011

— IBM Corp., 2013. IBM SPSS Statistics for Windows, Version 22.0.

— Johnson, E., Lasater, K., Hodson Carlton, K., Sideras, S., Siktberg, L., & Dillard, N.

— (2012). Geriatrics in simulation: role modeling and clinical judgment effect. Nursing Education Perspectives, 33(3), 176-180.

— Kim, J., Park, J.-H., Shin, S., 2016. Effectiveness of simulation-based nursing education depending on fidelity: a meta-analysis. BMC Med. Educ. 16, 152. doi:10.1186/s12909-016-0672-7

— Kim, S.-J., Kim, S., Kang, K.-A., Oh, J., Lee, M.-N., 2016. Development of a simulation evaluation tool for assessing nursing students' clinical judgment in caring for children with dehydration. Nurse Educ. Today 37, 45–52. doi:10.1016/j.nedt.2015.11.011

— Lasater, K., 2011. Clinical judgment: the last frontier for evaluation. Nurse Educ. Pract. 11, 86–92. doi:10.1016/j.nepr.2010.11.013

— Lasater, K., 2007. Clinical judgment development: using simulation to create an assessment rubric. J. Nurs. Educ. 46, 496–503.

— Lasater, K., Johnson, E. A., Ravert, P., & Rink, D. (2014). Role-modeling clinical judgment in an older adult simulation. Journal of Nursing Education, 53(5), 257-264.

— MacCallum, R. C. , Browne, M. W., & Sugawara, H. M. 1996.Power analysis and determination of sample size for covariance structure modeling. Psychological Methods, 1,130-149

— Martínez-Castillo, F., Matus-Miranda, R., 2015. Desarrollo de habilidades con simulación clínica de alta fidelidad. Perspectiva de los estudiantes de enfermería. Enferm. Univ. 12, 93–98. doi:10.1016/j.reu.2015.04.003

— National Academies of Sciences, Engineering, and Medicine., 2016. Assessing Progress on the Institute of Medicine Report The Future of Nursing. The National Academies Press, Washington, DC.

— Norman, J., 2012. Systematic review of the literature on simulation in nursing education. ABNF J. Off. J. Assoc. Black Nurs. Fac. High. Educ. Inc 23, 24–28.
— Oh, P.-J., Jeon, K.D., Koh, M.S., 2015. The effects of simulation-based learning using standardized patients in nursing students: A meta-analysis. Nurse Educ. Today 35, e6–e15. doi:10.1016/j.nedt.2015.01.019
— Seagull, F.J., Rooney, D.M., 2014. Filling a void: developing a standard subjective assessment tool for surgical simulation through focused review of current practices. Surgery 156, 718–722. doi:10.1016/j.surg.2014.04.048
— Shin, H., Park, C.G., Shim, K., 2015a. The Korean version of the Lasater Clinical Judgment Rubric: a validation study. Nurse Educ. Today 35, 68–72. doi:10.1016/j.nedt.2014.06.009
— Shin, H., Shim, K., Lee, Y., Quinn, L., 2014. Validation of a new assessment tool for a pediatric nursing simulation module. J. Nurs. Educ. 53, 623–629. doi:10.3928/01484834-20141023-04
— Shin, S., Park, J.-H., Kim, J.-H., 2015. Effectiveness of patient simulation in nursing education: meta-analysis. Nurse Educ. Today 35, 176–182. doi:10.1016/j.nedt.2014.09.009
— Shinnick, M.A., Woo, M.A., 2013. The effect of human patient simulation on critical thinking and its predictors in prelicensure nursing students. Nurse Educ. Today 33, 1062–1067. doi:10.1016/j.nedt.2012.04.004
— Tanner, C.A., 2006. Thinking like a nurse: a research-based model of clinical judgment in nursing. J. Nurs. Educ. 45, 204–211.

— Thompson, C., Aitken, L., Doran, D., Dowding, D., 2013. An agenda for clinical decision making and judgement in nursing research and education. Int. J. Nurs. Stud. 50, 1720–1726. doi:10.1016/j.ijnurstu.2013.05.003
— Thompson, C., Stapley, S., 2011. Do educational interventions improve nurses' clinical decision making and judgement? A systematic review. Int. J. Nurs. Stud. 48, 881–893. doi:10.1016/j.ijnurstu.2010.12.005
— Victor-Chmil, J., Larew, C., 2013a. Psychometric properties of the Lasater clinical judgment rubric. Int. J. Nurs. Educ. Scholarsh. 10. doi:10.1515/ijnes-2012-0030
— Wild, D., Grove, A., Martin, M., Eremenco, S., McElroy, S., Verjee-Lorenz, A., Erikson, P., ISPOR Task Force for Translation and Cultural Adaptation, 2005. Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: Report of the ISPOR Task Force for Translation and Cultural Adaptation. Value Health J. Int. Soc. Pharmacoeconomics Outcomes Res. 8, 94–104. doi:10.1111/j.1524-4733.2005.04054.x
— World Health Organization, 2009. Global standards for the initial education of professional nurses and midwives.

Tables and Figures:

Table 1 Scores for the LCJR evaluation items

	Simulation 1				Simulation 2			
	Mean	SD	Endorsement frequency (%)		Mean	SD	Endorsement frequency (%)	
			Lowest value	Highest value			Lowest value	Highest value
Focused observation	2.36	.792	12.6	6.9	2.50	.854	12.5	10.9
Recognizing deviations…	2.20	.790	17.2	3.4	2.38	.845	17.2	7.8
Information seeking	2.22	.769	16.1	5.7	2.33	.856	17.2	9.4
Prioritizing data	2.29	.806	20.7	4.6	2.42	.887	17.2	3.1
Making sense of data	2.14	.795	16.1	5.7	2.30	.790	17.2	4.7
Calm, confident manner	2.34	.819	14.9	5.7	2.33	.818	15.6	4.7
Clear communication	2.39	.812	24.1	5.7	2.36	.804	23.4	9.4
Well-planned intervention/ flexibility	2.09	.830	12.6	5.7	2.25	.926	6.3	9.4
Being skillful	2.28	.758	8.0	6.9	2.47	.755	10.9	3.1
Evaluation/self-analysis	2.55	.743	9.2	4.6	2.45	.733	4.7	3.1
Commitment to improvement	2.48	.729	12.6	6.9	2.59	.635	12.5	10.9

Table 2. Difference in the partial and total LCJR scores in the different assessment scenarios, differentiated by observers

	OSCE 1 Evaluator 1 Mean (SD)	**OSCE 2 Evaluator 1** Mean (SD)	**p**	**OSCE 1 Evaluator 2** Mean (SD)	**OSCE 2 Evaluator 2** Mean (SD)	**p**
Focused observation	2.39 (2.48)	2.48 (0.77)	0.575	2.3 (2.59)	2.59 (0.81)	0.131
Recognizing deviations…	2.22 (2.34)	2.34 (0.79)	0.372	2.13 (2.52)	2.52 (0.80)	0.044
Information seeking	2.17 (2.28)	2.28 (0.77)	0.615	2.26 (2.41)	2.41 (0.77)	0.462
Prioritizing data	2.39 (2.55)	2.55 (0.8)	0.398	2.13 (2.41)	2.41 (0.80)	0.152
Making sense of data	2.24 (2.24)	2.24 (0.74)	0.782	2.00 (2.45)	2.45 (0.83)	0.019
Calm, confident manner	2.39 (2.34)	2.34 (0.77)	0.796	2.26 (2.38)	2.38 (0.88)	0.549
Clear communication	2.41 (2.41)	2.41 (0.75)	0.915	2.33 (2.38)	2.38 (0.90)	0.760
Well-planned intervention/ flexibility	2.11 (2.34)	2.34 (0.71)	0.251	2.04 (2.24)	2.24 (0.94)	0.371
Being skillful	2.24 (2.55)	2.55 (0.67)	0.062	2.24 (2.55)	2.55 (0.85)	0.113
Evaluation/self-analysis	2.50 (2.59)	2.59 (0.70)	0.382	2.51 (2.41)	2.41 (0.82)	0.545
Commitment to improvement	2.48 (2.66)	2.66 (0.70)	0.232	2.40 (2.62)	2.62 (0.78)	0.221
TOTAL	**24.79 (26.79)**	**26.79 (6.87)**	**0.272**	**24.19 (26.97)**	**26.97 (7.99)**	**0.151**

Table 3. LCJR scores, differentiated by gender.

	Male (n=44) Mean (SD)	Female (n=108) Mean (SD)	p
Focused observation	2.41 (0.69)	2.42 (0.87)	0.918
Recognizing deviations…	2.18 (0.72)	2.31 (0.85)	0.517
Information seeking	2.20 (0.79)	2.29 (0.81)	0.664
Prioritizing data	2.41 (0.69)	2.32 (0.90)	0.463
Making sense of data	2.14 (0.67)	2.23 (0.84)	0.552
Calm, confident manner	2.30 (0.85)	2.36 (0.80)	0.618
Clear communication	2.34 (0.86)	2.39 (0.79)	0.706
Well-planned intervention/ flexibility	2.14 (1.00)	2.17 (0.82)	0.764
Being skillful	2.43 (0.87)	2.33 (0.71)	0.324
Evaluation/self-analysis	2.57 (0.76)	2.49 (0.73)	0.856
Commitment to improvement	2.48 (0.76)	2.55 (0.66)	0.470
TOTAL	25.59 (6.99)	25.56 (7.08)	0.888

Table. 4 Inter-rater correlations of individual dimensions

	OBSERVER 2										
OBSERVER 1	Focused observation	Recognizing deviations…	Information seeking	Prioritizing data	Making sense of data	Calm, confident manner	Clear communication	Well-planned intervention/ flexibility	Being skillful	Evaluation / self-analysis	Commitment to improvement
Focused observation	1.000										
Recognizing deviations…	.785	1.000									
Information seeking	.715	.699	1.000								
Prioritizing data	.689	.639	.579	1.000							
Making sense of data	.668	.679	.659	.738	1.000						
Calm, confident manner	.685	.698	.617	.613	.592	1.000					
Clear communication	.710	.700	.683	.620	.609	.845	1.000				
Well-planned intervention/ flexibility	.704	.684	.678	.624	.636	.717	.737	1.000			
Being	.584	.554	.601	.499	.488	.650	.657	.771	1.000		
Evaluation/self-analysis	.434	.489	.383	.339	.318	.410	.460	.441	.429	1.000	
Commitm	.417	.406	.340	.354	.335	.405	.422	.467	.47	.738	1.000

Table 5. Item-total correlation

	Item-total correlation	Squared multiple correlation	Alpha if item deleted
Focused observation	.799	.716	.928
Recognizing deviations…	.814	.731	.928
Information seeking	.750	.644	.930
Prioritizing data	.722	.637	.932
Making sense of data	.728	.660	.931
Calm, confident manner	.799	.759	.928
Clear communication	.827	.783	.927
Well-planned intervention/ flexibility	.813	.731	.928
Being skillful	.725	.644	.932
Evaluation/self-analysis	.534	.593	.939
Commitment to improvement	.521	.581	.939

Figure 1

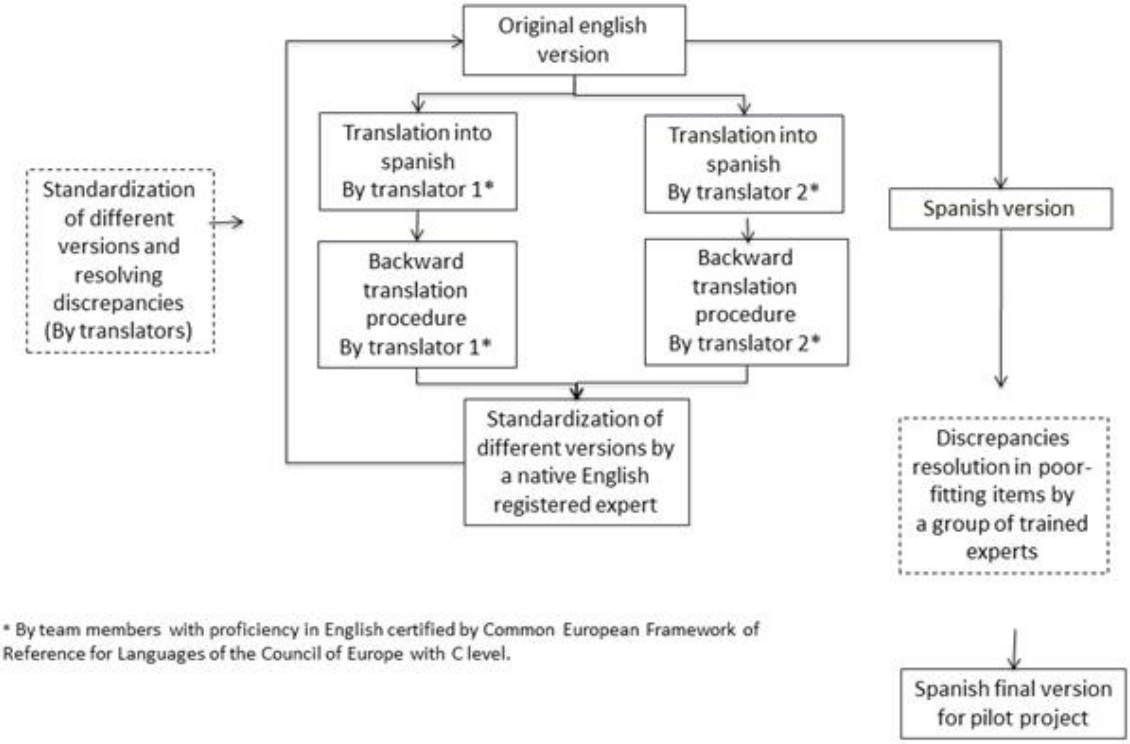
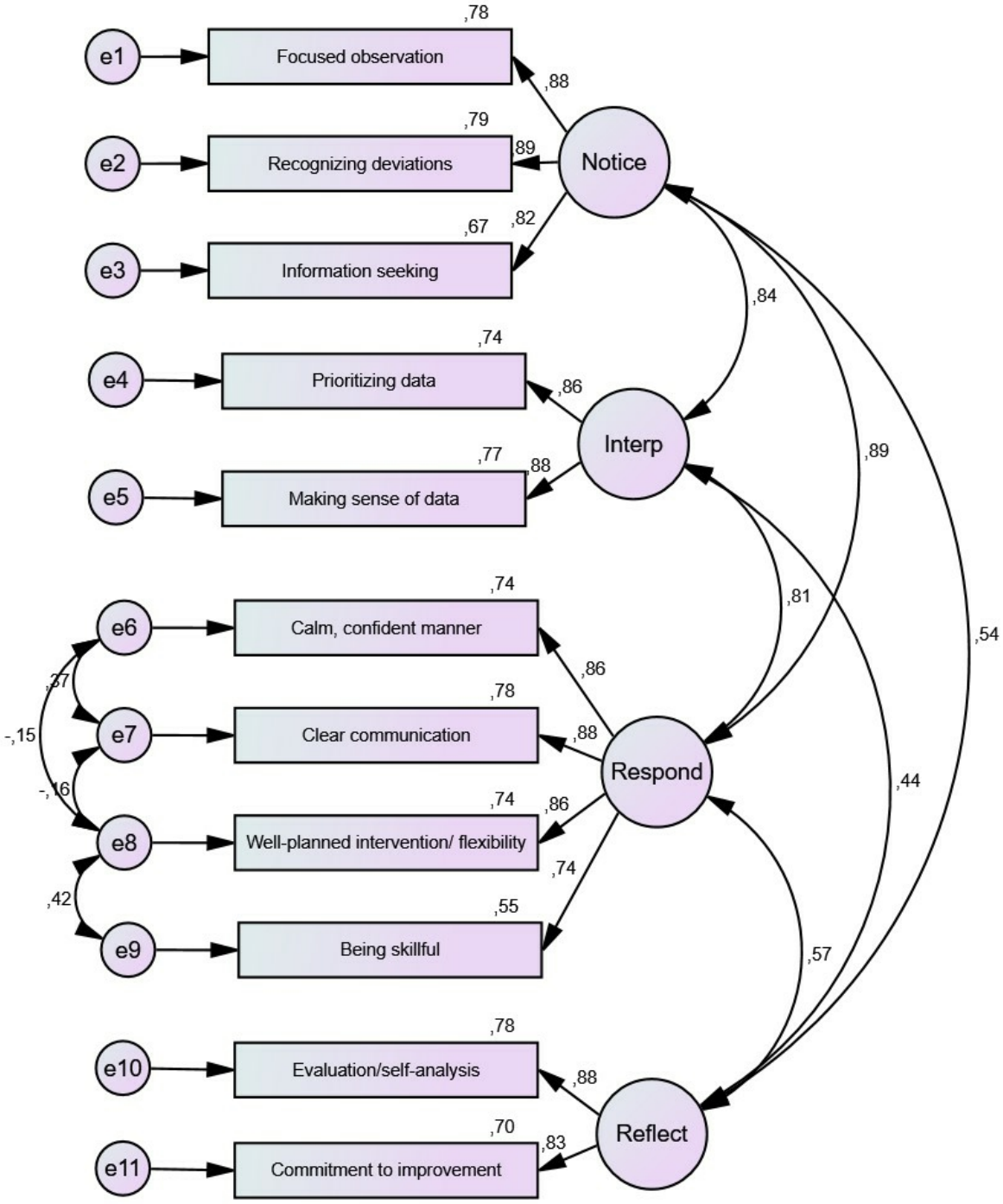
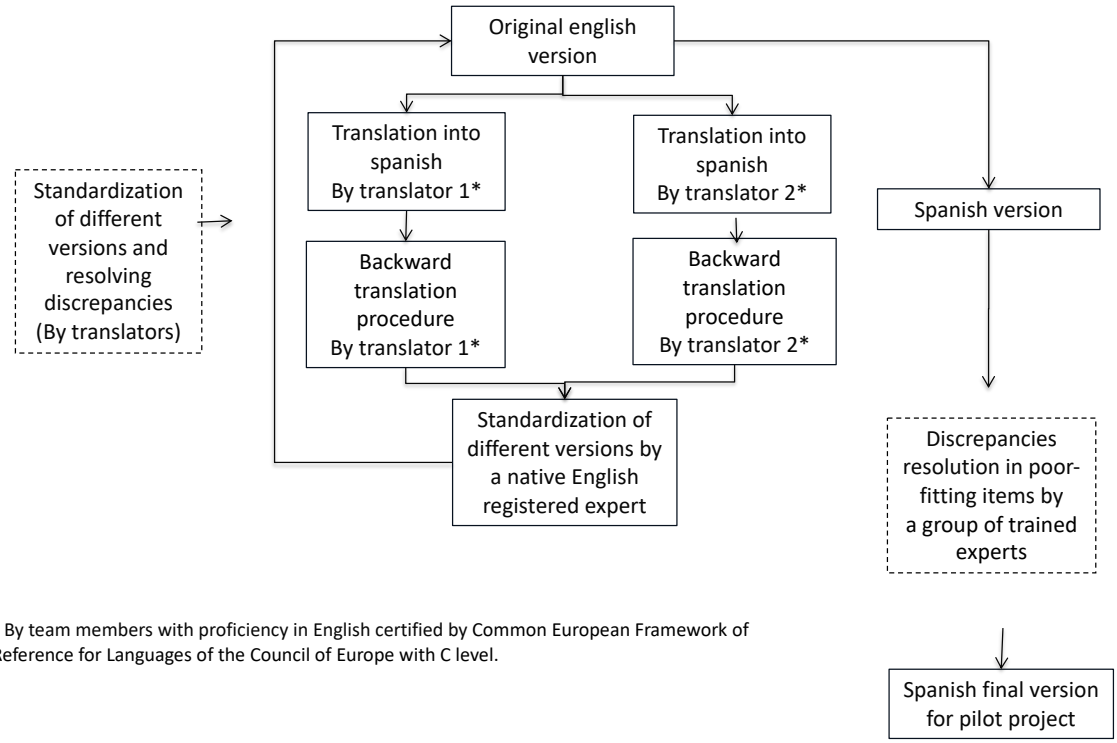

Original english version
Translation into spanish By translator 1*
Translation into spanish By translator 2*
Spanish version
Standardization of different versions and resolving discrepancies (By translators)
Backward translation procedure By translator 1*
Backward translation procedure By translator 2*
Standardization of different versions by a native English registered expert
Discrepancies resolution in poor-fitting items by a group of trained experts
* By team members with proficiency in English certified by Common European Framework of Reference for Languages of the Council of Europe with C level.
Spanish final version for pilot project

Figure 2


,78
e1
Focused observation
,88
,79
e2
Recognizing deviations
,89
Notice
,67
,82
e3
Information seeking
,84
,74
e4
Prioritizing data
,86
Interp
,77
,88
e5
Making sense of data
,89
,74
e6
Calm, confident manner
,86
,81
,54
,37
,78
-,15
e7
Clear communication
,88
Respond
-,16
,74
,86
e8
Well-planned intervention/ flexibility
,44
,74
,42
,55
e9
Being skillful
,57
,78
e10
Evaluation/self-analysis
,88
,70
,83
e11
Commitment to improvement
Reflect

Workflow



* By team members with proficiency in English certified by Common European Framework of Reference for Languages of the Council of Europe with C level.