International Journal of Management (IJM)
Volume 7, Issue 7, November–December 2016, pp.204–216, Article ID: IJM_07_07_021
Available online at
http://www.iaeme.com/ijm/issues.asp?JType=IJM&VType=7&IType=7
Journal Impact Factor (2016): 8.1920 (Calculated by GISI) www.jifactor.com
ISSN Print: 0976-6502 and ISSN Online: 0976-6510
© IAEME Publication

DESIGN, ADAPTATION AND CONTENT VALIDITY PROCESS OF A QUESTIONNAIRE: A CASE STUDY

M.E. José Nicolás Cardona Mora, Dr. Francisco Bribiescas Silva, Dr. Roberto Romero Lopez and Rosa Elba Corona Cortez
Universidad Autonoma de Ciudad Juarez, Mexico

ABSTRACT

One of the tools most used in the past decades for data collection is the questionnaire´s based surveys. In this article a theoretical review of the design, adaptation and content validity process of a questionnaire is performed. A case study of this process is presented to have a practical example of the of this research stage.

For the case study the objective is to create and develop a questionnaire up to the point of content validity. The final results gave a product of a questionnaire of 48 items and a content validity accepted with a value of a "V" of Aiken and accordance proportions indexes above of 0.80 in all the items.

Key words: Content Validity, V of Aiken, Questionnaire Design

Cite this Article: M.E. José Nicolás Cardona Mora, Dr. Francisco Bribiescas Silva, Dr. Roberto Romero Lopez and Rosa Elba Corona Cortez. Design, Adaptation and Content Validity Process of a Questionnaire: A Case Study. *International Journal of Management*, 7(7), 2016, pp. 204–216. http://www.iaeme.com/IJM/issues.asp?JType=IJM&VType=7&IType=7

1. INTRODUCTION

The human being is curious by nature and along his existence has developed different methods and ways to obtain the answer on the unknown things, such as observation and experimentation. These two methods gave birth to science and as Chalmers (1987) states "Science is based on what we can see, hear, touch, etc. Science is objective. Scientific knowledge is reliable knowledge because it is objectively proven knowledge".

The best way to create scientific knowledge is by an established methodology that will lead the researcher step by step thru the creation of this new knowledge.

Actually, for a for scientific research to be considered as such, it is necessary that a sequence of steps must be followed and not be altered, because doing so would jeopardize the validity and reliability of the study(Hernandez, Fernandez y Baptista, 1991). Within these steps, the researcher must collect field data that will validate their hypothesis.

Hernandez et al (1991) indicates that in order to collect data, the researcher must first select or design the measurement tool, and this tool must be previously validated and reliable. The selection of the tool is a very important step on any research, if is not the proper tool for the purpose of the research the data collected will be useless and valuable time will be lost, or even worst, the research will be invalidated.

In the scientific methodology exists various tools for data collection, but thanks to recent methodological developments, especially in statistical data analysis, survey methodology (based in questionnaires) has become one of the most common alternatives for researchers in recent decades (Meneses & Rodríguez, 2011).A survey is a technique that uses a set of standardized research procedures through which it collects and analyzes a range of data from a representative sample of cases of a population or universe, which aims to explore, describe, predict and/or explain a number of features" (Anguita, Labrador, & Campos, 2003)

Anguita et al (2003), states that the basic tool used in research by survey is the questionnaire, and it can be defined as "a document that collects in an organized way the indicators of the variables involved in the objective of the survey". Meneses & Rodríguez (2011) define a questionnaire as "a standardized instrument used to collect data during the field work in some quantitative researches, mainly of those carried out in surveys". Both definitions are very similar in their context, so, it can be said that a questionnaire is used as a standardize tool that allows the researcher to get the required information from the studied sample. And to avoid confusions, the difference between survey and questionnaire is that the survey is the complete process of data collection, meanwhile, the questionnaire is the physical tool with the questions that the participants will answer.

To design and/or adapt a questionnaire is not an easy task; it is not a simple set of questions that can be filled by anyone. For Meneses & Rodríguez (2011) a questionnaire must meet the following characteristics:

- Must produce quantitative data for statistical processing and analysis.
- Questions must be structured and directed to a particular group of people.
- Data must represent a given population.

These characteristics are very important and any questionnaire created for a scientific research must have them, but there is another characteristic that Meneses & Rodríguez (2011) did not mentioned, a questionnaire used for a scientific research must be "validated".

The validity process of a questionnaire is a very important step and it must be performed before the questionnaire is applied to the population´s sample. This validation will allow the researcher to have more certainty that his questions are understandable and will gather the information required for the research objective. This validation is part of the construction/adaptation process of a questionnaire; Carretero-Dios & Pérez (2007) indicate the following steps for this process:

- Study justification.
- Conceptual definition of the constructs to be evaluated.
- Construction and qualitative evaluation of the items.
- Statistical analysis of the items.
- Study of the dimensionality of the instrument (internal structure).
- Reliability estimation.
- Obtaining evidence of external validity.

It must be understood that to reach this part of the research process several steps must have been accomplished before, such as: problem definition, research objective, research viability, hypothesis, etc. For this article only the first three steps of the validation will be explained and developed within the case study.

2. STUDY JUSTIFICATION

In order to create or adapt a new data collection tool, the researcher must provide and justify the reasons of why is needed or what new data this tool will provide, and the conditions of the research that will allow to have a successful data collection. One of the first things to do is delimitate the characteristics of the questionnaire (Carretero-Dios & Pérez, 2007):

- *What will be evaluated:* To be able to answer this question the researcher must have done all the previous theoretical work of the research. To know what theories have influence on the research and how they will contribute to the formulation of the construct(s) will help the researcher to have more clarity at the moment of selecting the questions of the questionnaire.
- *Who will be evaluated:* This information must be identified since the start of the research. Depending on the population, the questions will be written in a certain way, because, it is not the same to apply a questionnaire to university students than industry professionals who are already working. Besides, it is also important to have the population well identified to be certain that the gathered information will be useful.
- *What is the purpose of the research:* The objective of the research must be clear at this stage of the investigation. The whole questionnaire will be design around the research objective and the questions will be oriented on obtaining the required information. Also, if the researcher has clear what is needed to be obtained from the questionnaire it will be easier the design of the questions.

3. CONCEPTUAL DEFINITION OF THE CONSTRUCTS TO BE EVALUATED

In order to create or adapt a questionnaire it is essential to have a correct definition of the construct(s) that will be evaluated, it´s important to indicate that some researches suffer from a proper conceptualization of the construct, which it is usually the result of poor literature review, and finally ends up affecting the quality of the questionnaire (Carretero-Dios & Pérez, 2007). One tool that can be used to have a better conceptual definition of the constructs is a “Specifications Table” that can include some of the following information: name of the construct, semantic definition, operational definition, the items or question related to the construct and the scale used. Is not only important operational and semantic definitions of the variables, it is also necessary to identify and define the relationship established between them and with other variables or “syntactic definition”(Muñiz, J., & Fonseca-Pedrero, E., 2009).

4. CONSTRUCTION AND QUALITATIVE EVALUATION OF THE ITEMS

The construction of the items or questions is one of the most important steps in the construction process the measuring instrument (Downing, 2006; Muñiz, J., & Fonseca-Pedrero, E., 2009; Schmeiser and Welch, 2006). The questions are the spine of the questionnaire, the main axis where all the data collected will converge. The information gathered will depend largely on how the questions are structured. Some recommendations for the construction of the questions are:they must be clear and simple, avoid technicalities, double negatives, or excessively detailed or ambiguous statements, personalization, be neutral, be one logic statement, perform simple calculations and be concise (Muñiz et al, 2005; Anguita et al, 2003).

To be able to develop a better quality on the items, an exhaustive literature review must be done by the researcher;investigate on the theoretical background and look for similar researches or questionnaires that content one or several of the constructs that will be evaluated. Hernandez et al (1991) mention that a common mistake made by a researcher is to produce an instrument to collect data without having reviewed the literature on the subject; and this inevitably leads to error or at least serious deficiencies in the questionnaire.

During this stage of the construction of the questionnaire, the researcher must perform a validity content process of the questionnaire.

5. VALIDITY CONTENT

The quality of a measuring instrument depends basically of two properties: reliability and validity, where the term reliability is usually used as synonym of repeatability, reproducibility or consistency; and validity refers to whether the procedure it is actually measuring the phenomenon it is wanted to be measured (Latour, Abraira, Cabello & Sánchez, 1997). This article will cover the part of "content validity" only, construct validity and criteria validity will be cover in future work.

According to Polit & Tatano (2006) content validity is define as "the degree to which a sample of items, taken together, constitute an adequate operational definition of a construct", and indicates that it involves two distinct phases: a priori efforts by the scale developer to enhance content validity through careful conceptualization and domain analysis prior to item generation, and a posteriori efforts to evaluate the relevance of the scale's content through expert assessment. For Carretero-Dios & Pérez (2007) content validity is the evidence that semantic definition was well contained in the formulated items; and for Latour et al (1997), content validity indicates the extent to which all the items that make up the index, covers the different areas or domains that wanted to be measured. So, in summary it can be said that content validity is a way to measure how much the items or questions in the questionnaire measure the construct that are being evaluated.

Escobar-Pérez & Cuervo-Martínez (2008) indicates that content validity is established in different situations, two of the most common are: a) design a test, and b) the validation of an instrument that was built for a different population, but was adapted by translation.

A common way to measure validity content is by the methodology of expert judge's review or expert's judgment (Wynd, Schmidt & Schaefer, M.,2003; Escurra, 1989; Blazquez, 2011; Merino y Livia, 2009; Mendoza & Garza, 2009; Dominguez & Villegas, 2012).

6. EXPERT JUDGES REVIEW METHODOLOGY

Currently the expert review is a widespread practice that requires interpreting and applying their results in a precise, efficient and with the entire methodological and statistical rigor way; and content validity is usually assessed through this practice (Escobar-Pérez & Cuervo-Martínez, 2008). Another definition is given by Wynd et al (2003), and he states that "the resulting instrument content validity is based mainly on the judgment, logic, and reasoning of the researcher with validation from a panel of judges holding expertise in the domain of content".

Concerning the appropriate number of judges to consult, there is not a fixed or established number, in fact a wide variety of criteria between authors has been found. For example, Escobar-Pérez & Cuervo-Martinez (2008) indicates that "the number of judges to be used at a review depends on the level of expertise and diversity of knowledge; however, the decision about how many experts is appropriate varies among authors" and mention examples in a range from two to twenty. Wynd et al (2003) warns that an increased number of experts (raters, observers, or judges) and a larger number of categories for data assignment yield greater absolute agreement and increase the risk of chance agreement.

As part of the scientific method, the expert review of a questionnaire to estimate the validity content is a process that must be systematic, and Escobar-Pérez & Cuervo-Martinez (2008) suggest a step by step guide to implement this process in an efficient and organized way.

Steps

- *Define the objective of the expert review process.* The researcher must have clear what is the purpose of the expert review, if it is due a translation, or a cultural adaptation to use the questionnaire in another country; or to validate the creation of a new measuring tool.
- *Experts or Judges Selection.*Pedrosa, I., Suárez-Álvarez y García-Cueto, E. (2013) indicate that "the appropriate selection of experts is a critical issue at the time of establishing this type of validity". The researcher must establish the criteria that the experts must accomplish, for example in the article of Mendoza & Garza (2009) the authors established that the experts must be an academic-practical research

methodology expert (with a postgraduate level or be active researcher) and/or business expert in the field of organizational innovation (working in a department of innovation and development, with a minimum seniority of three years and with a manager or director position).

- *Clarify the dimensions; the indicators that are measuring each of the test items.* It is important to clarify this information and do not assume that the judge will understand what it is supposed to evaluate.
- *Specify the objective of the questionnaire.* The judges or experts must know the purpose of the questionnaire that will allow understanding the context and providing a better evaluation.
- *Set the differential weights of the dimensions of the test.* This is done only when the test have different weights for some dimensions.
- *Design the Expert´s review format.* There is no specific format to give to the experts to perform their evaluation, the research must choose the design and type of spreadsheet for this step, but it is important that must be easy to understand
- *Calculate the accordance among the experts.* The researcher must understand the different types of indexes to calculate the accordance among experts, and have the expertise to perform all the statistical process.
- *Elaborate conclusions.* The researcher must be capable to interpret the results of the accordance test, understand what the numbers are saying, take decisions of what to do next and elaborate the results summary.

7. TYPES OF CONTENT VALIDITY ASSESSMENTS

Exist several indexes or assessments for content validity, each of them have their own particular characteristics, and a reason of why they were created. Below will review some of these assessments tools:

- *Proportion agreement* – The Proportion agreement procedure as a Content Validity Index (CVI) is explained by Wynd et al (2003) as: "allow two or more raters to independently review and evaluate the relevance of a sample of items to the domain of content represented in an instrument". The main concern with the CVI is that it is an index of inter-rater agreement that simply expresses the proportion of agreement and agreement can be inflated by chance factors (Cohen, 1960; Polit & Tatano, 2006; Wynd et al, 2003)
- *Kappa index* – After the CVI and due to the concerns about proportion agreement, the Kappa coefficient was proposed, which quickly becomes the index most used in biological and social sciences is (Escobar-Pérez & Cuervo-Martinez, 2008). The kappa index represents the proportion of agreement remaining after chance agreement is removed. But there are two inconvenient with the Kappa index, 1) the value does not give any indication of precision, in other words, it does not give any information of the variability it may have, and there is not possible to make any statistical inference, 2) Kappa index does not give any information about the quality of the measurement performed by the observers, as it is designed only to estimate the magnitude of the correlation between the two (Cerda & Villarroel 2008).
- *Aiken index* – Sireci & Geisinger (1995) indicate that "the statistical significance of the Aiken index provides a practical measure of Subject-matter experts (SME) congruence. The Aiken index and averaged relevance ratings provided similar information; therefore, computing both indexes is probably not necessary. The Aiken index appears preferable because it can be evaluated for statistical significance" With the last statement it is understandable why Aiken index can be more useful than proportion agreement and kappa index, besides, the Aiken index can be use with ratings dichotomous (values 0 or 1) or polytomous (values from 1 to 5); the calculation is simple and objective and as it has been stated previously, it is statistical reliable (Escurra, 1989).

The formula used to calculate the V of Aiken index is (Merino and Livia, 2009):

$$V = \frac{\bar{X} - l}{k}$$

Where:

V=accordance level index

$\bar{X}$=average of the experts grades

l=minimum grade possible

k=range from the likert scale.

This index can be complemented with the calculation of the confidence intervals by the "score" method, and according to Merino and Livia (2009) it has very good properties for the analysis because it does not depend on the normal distribution of the variable; it is asymmetrical with respect to the variable and is highly accurate. The formulas to calculate these intervals are:

$$L = \frac{2nkV + z^2 - z\sqrt{4nkV(1-V) + z^2}}{2(nk + z^2)}$$

$$U = \frac{2nkV + z^2 + z\sqrt{4nkV(1-V) + z^2}}{2(nk + z^2)}$$

Where:

L = Lower interval limit

U = Upper interval limit

Z = Normal standard distribution value

V = Aiken index

N = number of experts judges.

The criteria used to decided accept or reject an item based on the expert review with the V of Aiken will depend on the researcher and the type of study, for example, Merino and Livia (2009) indicate that a value higher of 0.50 are items that must be kept in the instrument but it is low criteria, meanwhile for a more conservative standpoint must be higher than 0.70, same value accepted by Sireci & Geisinger (1995).

8. CASE STUDY

The case study presented in this article is the design and content validity of the questionnaire that will be used for the doctoral thesis "Administrative model for deployment and implementation of Continuous Improvement into production processes in the manufacturing sector in Ciudad Juarez, Chihuahua". Up to this pointall the previous research stages have been completed, and the information will be used in this article to explain and support the decisions during the creation of the questionnaire.

The methodology used for the design and validation of the questionnaire is based (but not strictly followed) in the first 3 steps proposed by Carretero-Dios & Pérez (2007), explained previously in this article, which they are: a)Study justification, b) Conceptual definition of the constructs to be evaluated, and c) Construction and qualitative evaluation of the items.

9. STUDY JUSTIFICATION

The objective of the research in this case study is to create an administrative model to help the deployment and implementation of Continuous Improvement (CI) in manufacturing plants. In order to be able to design the model, it is necessary to understand what critical factors are important for the CI implementation agents in the manufacturing sites; and how is the current condition of those critical factors inside those plants. To get this information it was decided that a field survey is the best option to collect the required data, but first is necessary to have the variables that will be measured.

Carretero-Dios & Pérez (2007) suggest in his procedure that it is important to delimitate the following characteristics of the questionnaire.

- *What will be evaluated:* It will be evaluated the level of importance of the variables identified as necessary (based on the literature review) for a good CI implementation, and the current situation in the manufacturing sites regarding to these variables.
- *Who will be evaluated:* The population that is beingstudied is the "maquiladoras" (name for manufacturing plants in Mexico) that are dedicated to the automotive industry in Juarez, Chihuahua, Mexico. The population is well identified; there are about a total of 70 automotive maquiladoras in Juarez. Inside the plants, the appropriate people to answer the questionnaire must be the direct responsible for CI implementation and must have a position of supervisor, manager (department or plant) and/or director. The role of the respondents may vary for each company, some of them have an independent CI department, or in some others this implementation is responsibility of the engineering or quality departments.
- *What is the purpose of the research:* The objective of the research is to design an administrative model that will help the manufacturing sites to have better management of their CI resources, or will facilitate the implementation and deployment of CI. The questionnaire must provide an assessment of which variables must be included in the model and if there is a tendency to some of the studied variables.

10. CONCEPTUAL DEFINITION OF THE CONSTRUCTS TO BE EVALUATED

It's indispensable to have a good understanding of the conceptual definitions of the variables or constructs that will be evaluated in order to have a better design of the items. During the stages of theoretical and reference framework the constructs to be measure were identified through the following process.

One of the previous stages in the research was to identify the variables that will be measured and an extensive literature review was done in books, journals and doctoral thesis from different online data bases. The main variables are "Continuous Improvement" (CI) and "Productive Processes" (PP). The literature review showed that CI variable is too ambiguous and it needs to be broken in sub-variables that are easier to measure. Most of the reviewed articles showed that the most common sub-variables in CI are: management commitment, leadership, training, collect employees' ideas, strategy, CI implementation, employees' recognition and communication; on the other hand, PP is a tangible variable to be measure that are related with the productive system, such as: scrap, efficiency or productivity, absenteeism, turnover, etc.

Having the characteristics clear, the next step and first option, was to do an extensive search in the literature for validated questionnaires that covers the variables identified in the research and that can be adapted to the environment of the current investigation. The search gave negative results, it wasn´t found the measuring instrument that covers the characteristics needed, so the adaptation of a validated instrument has to be discarded.

Due that the first option was not possible, now the measuring instrument needs to be design and created. But the questionnaire was not created from zero, instead what was done, was to search for questionnaires containing the variables studied and take those items as a basis to create new items that suit the environment of the research. With this in mind, eight questionnaires were selected from the following authors: Katts (2013); Vahed (2012); Jaca, Tanco, Santos, Mateo & Viles (2010); Yarto (2012); Madrigal (2012); Gadea (2006) and Robert, Probst, Martocchio, Drasgow & Lawler (2000).

The conceptual definitions for the variablesand how they are related to CI are as follows:

- *Management Commitment:* Management commitment is critical for any CI activity; it gives the program a full recognition (Radnor et al 2006).It is important to mention that this commitment should not be present only on at the beginning of the process, but in a continuous manner through follow up activities.
- *Leadership:* Dubrin (1990) define leadership as "the process of influencing the activities of a individual or group to achieve certain objectives in a given situation". Leadership has a huge impact in CI, the success or failure of an implementation depends directly from having a strong or weak leadership.
- *Employees' ideas Collection:* Collecting employees' ideas is a good way to motivate and enhance creativity, that if it is given the proper follow up, can became in powerful innovations; the most common

way to do it is by the implementation of a suggestion system. Ekvall (1971) define a suggestion system as: and administrative procedure for collection, judging and compensating ideas, which are conceived by employees of the organization.

- *Strategy:* Strategy can be defined as a pattern in a stream of decisions and actions that imposes stability in an organization (Mintzberg, 1987). CI strategy will guide the core team thru the implementation with the objective of having a clear path to success.
- *CI implementation:* Dynamic process that involves people, strategy and knowledge to establish procedures, tools and methodologies that are focused in enhance the output of a process.
- *Employees' recognition:* Process of award an employee for a valuable contribution to the organization, in CI implementation is very important to give credit to the employees who are contributing to improve the company; this keeps high the morale of the team.
- *Communication:* Communication can be defined as the process of sharing information between individuals or groups, and it can in different ways (Boon et al, 2006).Communication is vital for any type of CI implementation program, as Radnor et al (2006) states that "Effective communication is important to ensure Lean is successfully implemented".
- *Training :* Training and development is the process pf providing people with the necessary skills and competencies to perform better their jobs, and it has been recognized as essential to the implementation of continuous improvement environment (Vahed, 2012).

11. CONSTRUCTION AND QUALITATIVE EVALUATION OF THE ITEMS

Having the characteristics clear, the next step and first option, was to do an extensive search in the literature for validated questionnaires that covers the variables identified in the research and that can be adapted to the environment of the current investigation. The search gave negative results, it wasn´t found the measuring instrument that covers the characteristics needed, so the adaptation of a validated instrument has to be discharged.

Due that the first option was not possible, now the measuring instrument needs to be designed and created. But the questionnaire was not going to be created from zero, what was done, was to search for questionnaires containing the variables studied and take them as a basis to create new items that suit the environment of the research. With this in mind, seven questionnaires were selected from the following authors: Katts (2013); Vahed (2012); Jaca, Tanco, Santos, Mateo & Viles (2010); Yarto (2012); Madrigal (2012); Gadea (2006) and Robert, Probst, Martocchio, Drasgow & Lawler (2000).

Based on the description of the items in the sevenquestionnaires from the authors mentioned above, a selection of the questions that most adapt to the research were defined (123 questions), from there, questions that have the same meaning or low relevance for the research were filtered until the questionnaire reached the desired size of 52 questions.

With the first draft of the questionnaire ready, it is time to start the validation process and the first step is to perform a "Content Validity" test.

12. CONTENT VALIDITY TEST

The test chosen for the content validity is by "Expert Judge Review" methodology, and it was followed the procedure of eight steps proposed by Escobar-Pérez & Cuervo-Martinez (2008) that is described previously in this article.

Steps

- *Define the objective of the expert review process:* The objective of the experts review is to identify and evaluate the items from a new questionnaire. Two assessments must be evaluated in each item, a) How much the item belongs to the construct where it is included, and b) How clear is the question to be understood.

- *Experts or Judges Selection:* The judges must be a Continuous Improvement expert, have experience on CI for at least 5 years and be currently working on a CI position (supervisor, manager, consultant). A total of 7 experts were selected according the established requirements.
- *Clarify the constructs:* The constructs definitions were explained to the experts when the review format was delivered to them. A brief discussion of each construct was part of the explanations, the same with the instructions of the review format.
- *Specify the objective of the questionnaire:* On the same explanation of step #3 it was explained the purpose of the research and what are the objectives, the experts did mentioned that this explanation help them to understand better at the time of doing the review.
- *Set the differential weights of the dimensions of the test:* The questionnaire has no different weights for the constructs.
- *Design the Expert´s review format:* The review format was created with the advice and recommendations of a professor from the Management Doctoral program that is an expert on questionnaires designs. Each question was evaluated by the experts on two assessments (construct belonging and clarity) and quantify them in a likert scale of five categories from 1 to 5.
- The review format can be provided if requested.
- *Calculate the accordance among the experts:* The results of the experts review was measured by the V of Aiken assessment index, and complemented with the confidence intervals to add more certainty to the study, the confidence intervals were calculated at a 95% of confidence. It was decided to not use the index Kappa of Cohen due that this index suits better for 2X2 tables (2 experts, 2 responses), decision based on Merino & Livia (2009). Also as it was previously stated, kappa does not provide information about the quality of the item.

  The minimum value for the V of Aiken will be of 0.70 based on the conservative criteria mentioned by Merino & Livia (2009). With the confidence intervals the value that will have to be higher of 0.70 to accept if the item is rejected or accepted is going to be the lower limit.

  In order to have more certainty on the accordance value, it was also calculated the agreement proportions, the average accordance proportion and the expected accordance proportion (EAP). Also was calculated the following values: sum of scores, average, standard deviation, minimum score and maximum score.
- *Results and conclusions for case study:* On the first evaluation for the assessment of "How much the item belongs to the construct where it is included" only the first 48 items were evaluates due the last four are additional questions that will provide information that is not included in any of the constructs.

The results of the first evaluation showed that eight out of nine constructs exceed the minimum value of 0.70 for V of Aiken in the Lower Confidence Limit (LCL), and 0.80 for the accordance proportions; the only construct that had a low score was "Productive Systems". Refer to table 1 for the results.

As it can be seen in table 1, the value of the lower confidence limit in the construct "Production System" is below 0.70 (marked in bold), this value indicate that the items in the construct should be revised to understand which items are the problem and consider to revise them or eliminate them.

In table 2 it can be seen the results of each expert to each item for this construct and the results. The questions Q3, Q4 and Q5 (marked in bold) are the ones that are reducing the score of the construct as a unit. Experts 1, 2 and 4 were consulted regarding those items and why they gave such a low score, all of them agreed that on their opinion those items in particular are not relevant for the construct.

Table 1 Results of the first evaluation for assessment in Belonging

Construct	V of Aiken	Lower Confidence Limit	Expected Accordance Proportion	Average Accordance Proportion
Commitment	0.923	0.799	92%	94%
Strategy	0.951	0.792	98%	96%
Leadership	0.911	0.738	93%	93%
Production System	0.857	**0.673**	80%	89%
Communication	0.920	0.750	96%	94%
Ideas	0.988	0.847	100%	99%
Training	0.964	0.811	100%	97%
CI Implementation	0.982	0.838	100%	99%
Recognition	1.000	0.868	100%	100%
Total	0.936	0.772	95%	95%

Source: Own Elaboration

After consulting the experts, the complete results were analyzed again, item by item, and it was found one more item that have a lower score for the LCL and in the EAP, item #5 from the construct "commitment". Same procedure of consulting the experts was done (but in this case was with experts 5 and 6), and the opinions were the same that last time. So, after this illustrative exercise of consulting the experts in the items with low scores it was decided to eliminate the four items, and then run a second evaluation for the questionnaire, but now with only 44 items.

In table 3 are the results of the new assessment for "How much the item belongs to the construct where it is included", now all nine constructs have values above of 0.70 for V of Aiken and the LCL, and above of 0.80 for the accordance proportions. An increase in the scores are noticeable for construct "Productive Systems", V of Aiken increase 0.063, the LCL increase 0.077 and the EAP increase 13%.

The results for the assessment of "Clarity was much better, but first is necessary to say that in these results are included the 52 items due that they belong to the first assessment of the questionnaire. In table 4 are the results of the evaluation.

Table 2 Results of first evaluation for assessment in Belonging for construct "Productive Systems"

Expert	Q1	Q2	Q3	Q4	Q5	Q6	Q7
1	5	3	3	3	3	5	4
2	5	4	4	3	3	5	5
3	5	5	5	5	5	5	5
4	5	5	3	3	3	4	4
5	5	5	5	5	5	5	5
6	5	5	5	5	5	5	5
7	5	4	4	5	4	5	3
V of Aiken	1.00	0.86	0.79	0.79	0.75	0.96	0.86
Lower Confidence Limit	0.91	0.72	0.64	0.64	0.60	0.85	0.72
Expected Accordance Proportion	100%	86%	71%	57%	57%	100%	86%
Average Accordance Proportion	100%	89%	83%	83%	80%	97%	89%

Source: Own Elaboration

Table 3 Results of the second evaluation for assessment in "Belonging"

Construct	V of Aiken	Lower Confidence Limit	ExpectedAccordance Proportion	Average Accordance Proportion
Commitment	0.940	0.821	95%	95%
Strategy	0.951	0.792	98%	96%
Leadership	0.911	0.738	93%	93%
Production System	0.920	0.750	93%	94%
Communication	0.920	0.750	96%	94%
Ideas	0.988	0.847	100%	99%
Training	0.964	0.811	100%	97%
CI Implementation	0.982	0.838	100%	99%
Recognition	1.000	0.868	100%	100%
Total	0.936	0.772	89%	95%

Source: Own Elaboration

The final results give the confidence that the content of the items is validated and the questionnaire can be applied on a first small sample to do the construct and criteria validity.

Table 4 Results of the second evaluation for assessment in "Clarity"

Construct	V of Aiken	Lower Confidence Limit	Expected Accordance Proportion	Average Accordance Proportion
Commitment	0.929	0.806	94%	94%
Strategy	0.942	0.823	98%	95%
Leadership	0.911	0.782	89%	93%
Prod. System	0.929	0.806	90%	94%
Communication	0.955	0.842	100%	96%
Ideas	0.940	0.821	95%	95%
Training	0.938	0.817	95%	95%
CI Implementation	0.964	0.855	100%	97%
Recognition	0.964	0.855	100%	97%
Additional	0.921	0.796	100%	94%
Total	0.937	0.816	96%	95%

Source: Own Elaboration

13. CONCLUSION

The design or adaptation of a questionnaire is not an easy task, it requires a lot of work and investigation, first on the literature review to have the certainty that the items placed on the questionnaire have solid foundations, and then in the validation test it must be performed to the measuring instrument. In this paper it was only presented to the step of content validity and it is only one third of the validation test that must be performed to the questionnaire before it can be applied to a representative sample.

Even that it is a lot of work, the validation process of a questionnaire is a step that cannot be skipped; the information that this process provides to the research is very valuable and gives the researcher another tool to prove that the investigation is being done properly.

Future Steps

The next steps for this research is to apply the questionnaire to a small sample to perform the construct and criteria validity tests and complete the validation of the measuring instrument before it is apply to a representative sample of the population that its being studied.

REFERENCE

[1] Anguita, J. C., Labrador, J. R., & Campos, J. D. (2003). La encuesta como técnica de investigación. Elaboración de cuestionarios y tratamiento estadístico de los datos (I). *Atención primaria*, 31(8), 527-538

[2] Blazquez, A. (2011). Diseño y validación de un cuestionario para analizar la calidad en empleados de servicios deportivos públicos de las mancomunidades de municipios extremeñas. *E-balonmano. com: Revista de Ciencias del Deporte*, 7(3), 181-192.

[3] Boon, K., Safa, M. & Arumugam, V. (2006). TQM practices and affective commitment: A case of Malaysian semiconductor packaging organizations. *International journal of management and entrepreneurship*, 27(1), 37-55.

[4] Carretero-Dios, H., & Pérez, C. (2007). Normas para el desarrollo y revisión de estudios instrumentales: consideraciones sobre la selección de tests en la investigación psicológica. *International Journal of Clinical and Health Psychology*, 7(3), 863-882.

[5] Cerda, J. & Villarroel del P., L. (2008). Evaluación de la concordancia interobservador en investigación pediátrica: Coeficiente de Kappa. *Revista Chilena de Pediatría*, 79(1), 54-58.

[6] Chalmers, A. (1987). *What Is This Thing Called Science?*Philadelphia: Open University Press.

[7] Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37-46.

[8] Dominguez, S., & Villegas, G. (2012). Estimación de la validez de contenido de una escala de calidad de vida para personas adultas con discapacidad intelectual. *Revista de Psicología de Arequipa*, 2(1), 207-219.

[9] Downing, S. M., & Haladyna, T. M. (2006). *Handbook of test development*. Lawrence Erlbaum Associates Publishers

[10] Dubrin, A. J. (1990). *Essentials of management*. Cincinnati, OH: South-Western Publishing.

[11] Ekvall,G. (1971). *Creativity at the Place of Work*. Stockholm: Reklamlito.

[12] Escobar-Pérez, J., & Cuervo-Martínez, A. (2008). Validez de contenido y juicio de expertos: una aproximación a su utilización. *Avances en medición*, 6, 27-36.

[13] Escurra, L. (1989) Cuantificación de la Validez de Contenido por Criterio de Jueces. *Revista de Psicología - PUCP*, 6, 103-111.

[14] Gadea, A. R. (2006). *Factores que facilitan el éxito y la continuidad de los equipos de mejora en la empresas industriales. Modelo de implantación, aplicación y medición de los resultados en una empresa piloto*. (Doctoral dissertation) Universitat Politècnica de Catalunya.

[15] Hernandez, R., Fernandez, C. y Baptista, P. (1991). *Metodología de la Investigación*. Naucalpan, Edo. de México, México: McGraw-Hill Interamericana de México S.A. de C.V.

[16] Jaca, C., Tanco, M., Santos, J., Mateo, R., & Viles, E. (2010). Sostenibilidad de los sistemas de mejora continua. *Intangible Capital*, 6(1), 51-77

[17] Katts, R. (2013). *The Sustainability of Continuous Improvement (CI) Initiatives in an Original Equipment Manufacturer (OEM) Paint Shop Environment* (Doctoral dissertation), Nelson Mandela Metropolitan University

[18] Latour, J., Abraira, V., Cabello, J. B., & Sánchez, J. L. (1997). Las mediciones clínicas en cardiología: validez y errores de medición. *Revista Española de Cardiología*, 50(2), 117-128.

[19] Madrigal, J. (2012). *Assessing Sustainability of the Continuous Improvement Through the Identification of Enabling and Inhibiting Factors* (Doctoral dissertation) Virginia Polytechnic Institute and State University.

[20] Mendoza, J., & Garza, J. B. (2009). La medición en el proceso de investigación científica: Evaluación de validez de contenido y confiabilidad. *Innovaciones de Negocios*, 6(1), 17-32.

[21] Meneses, J. and Rodríguez, D. (2011). El cuestionario y la entrevista. Retrieved from http://femrecerca.cat/meneses/files/pid_00174026.pdf

[22] Merino Soto, C. & Livia Segovia, J. (2009). Intervalos de confianza asimétricos para el índice de validez de contenido: Un programa Visual Basic para la V Aiken. *Anales de Psicología*, 25, 169-171.

[23] Mintzberg, H. (1987). *Crafting strategy*. Boston, MA: Harvard Business School Press.

[24] Muñiz, J., Fidalgo, A. M., García-Cueto, E., Martínez, R., y Moreno, R. (2005*). Análisis de los ítems*. Madrid: La Muralla.

[25] Fahd N. Al-Wesabi, Adnan Z. Alsakaf And Kulkarni U. Vasantrao. A Zero Text Watermarking Algorithm Based on the Probabilistic Patterns for Content Authentication of Text Documents. *International Journal of Computer Engineering and Technology (IJCET)*, 4 (1), 2013, pp. 284-300.

[26] Muñiz, J., & Fonseca-Pedrero, E. (2009). Construcción de instrumentos de medida para la evaluación universitaria. *Revista de Investigación en Educación*, 5, 13-25

[27] Pedrosa, I., Suárez-Álvarez y García-Cueto, E. (2013). Evidencias sobre la Validez de Contenido: Avances Teóricos y Métodos para su Estimación. *Acción Psicológica*, 10(2), 3-20.

[28] Polit, D., & Tatano, C. (2006). The content validity index: are you sure you know what's being reported? Critique and recommendations. *Research in Nursing and Health*, 29(5), 489-497.

[29] Radnor, Z., Walley, P., Stephens, A., & Bucci, G. (2006). *Evaluation of the Lean Approach to Business Management and It's Use in the Public Sector*. Edinburgh: Scottish Executive Social Research. Recuperado de http://www.scotland.gov.uk/Resource/Doc/129627/0030899.pdf

[30] Robert, C., Probst, T. M., Martocchio, J. J., Drasgow, F., & Lawler, J. J. (2000). Empowerment and continuous improvement in the United States, Mexico, Poland, and India: predicting fit on the basis of the dimensions of power distance and individualism. Journal of Applied Psychology, 85(5), 643.

[31] Roshani Pasalkar, Raj Makwana and Prof.Dr. S. D. Joshi. Image and Annotation Retrieval Via Image Contents and Tags. *International Journal of Computer Engineering and Technology (IJCET),* 6 (6), 2015,pp. 51-63.

[32] Sánchez-Meca, J. (2010). Como realizar una revisión sistemática y un meta-analisis. *Aula Abierta,* 38(2), 53-64.

[33] Schmeiser, C. B., & Welch, C. J. (2006). Test development. *Educational measurement*, 4, 307-353.

[34] Sireci, S. & Geisinger, K. (1995). Using Subject Matter Experts to Assess Content Representation: An MDS Analysis. *Applied psychological measurement*, 19(3), 241-255.

[35] Vahed Prevashini, P. (2012). *Continuous improvement and employee attitudes in a manufacturing concern* (Doctoral dissertation) North-West University.

[36] Wynd, C. A., Schmidt, B., & Schaefer, M. A. (2003). Two quantitative approaches for estimating content validity. *Western Journal of Nursing Research*, 25, 508–518.

[37] Yarto , M. (2012). *Modelo de Mejora Continua en la Productividad de Empresas de Cartón Corrugado del Área Metropolitana de la Ciudad de México* (Doctoral dissertation) Instituto Politécnico Nacional.