Consistency analysis from a school ethnography: approach to a validation of large-scale application questionnaires

Análisis de consistencia a partir de una etnografía escolar: aproximación a una validación para cuestionarios de gran escala

Magdalena Sánchez, Pedro Moreno, Angélica Vicuña

Dirección Evaluación y Estudios, Estudios y Consultorías Focus

Abstract

This article presents the research carried out for Agencia de Calidad to investigate the validity of the instruments used in 2014 to measure the Personal and Social Development standards in students, parents/guardians and teachers in sixth grade. Specifically, it presents what has been done to fulfill one of the specific objectives of the research, having as a starting point a consistency analysis between the quantitative results of the SIMCE instrument and the results of an ethnographic development, carried out in 18 schools along the country. The importance of this study relies in the comprehension and contextualization of possible adjustments to the contents of the questionnaires and it also gives observations about the scope and boundaries of the use of large scale instruments to measure Personal and Social Development Indicators. It also describes a methodology that allows integrating the ethnography to the content validity analysis of a group of instruments that try to measure non- academic standards from the consulted people's perception. The results, taken form the triangulation of quantitative and qualitative techniques, show medium-high consistency between the results of the two sources of information. It seems interesting to enhance this kind of methodology for consistency analysis to validate instruments of large scale evaluation.

Keywords: large scale educational evaluation, consistency in-between methodologies, non-academic indicators, educational quality

Resumen

El presente artículo da cuenta de una investigación realizada para la Agencia de Calidad de la Educación en la que se indaga en la validez de los instrumentos utilizados el 2014 para la medición de los Indicadores de Desarrollo Personal y Social en estudiantes, apoderados y profesores de sexto básico. Específicamente, se presenta lo realizado para el cumplimiento de uno de los objetivos específicos de la investigación, a partir de un análisis de consistencia entre los resultados cuantitativos del instrumento Simce y los resultados de un desarrollo etnográfico, realizado en 18 establecimientos educacionales del país. La relevancia del estudio radica en la contextualización y comprensión de posibles adecuaciones a los contenidos del cuestionario y entrega observaciones acerca de los alcances y límites del uso de instrumentos de gran escala para medir los IDPS. Además, se describe una metodología que permite integrar la etnografía al análisis de validez de contenido de un conjunto de instrumentos que buscan medir indicadores no académicos desde la percepción de los propios actores escolares. Los resultados, mediante la triangulación de técnicas cuantitativas y cualitativas, dan cuenta de un grado de consistencia medio-alto entre los resultados de ambas fuentes de información. Resulta interesante relevar esta metodología de análisis de consistencia para validar instrumentos de evaluación a gran escala.

Palabras claves: evaluación escolar a gran escala, consistencia entre metodologías, indicadores de Desarrollo Personal y Social, calidad educativa

Considering the importance of non-academic indicators in the development of students and the relevance they take on in the school context, the Agencia de Calidad de la Educación (Education Quality Agency) commissioned a study that among other objectives, would allow evaluation of the degree of consistency between the results from questionnaires on Education Quality and Context and what really takes place in schools in terms of dynamics and the perceptions of the stakeholders. This would enable the determination of whether the responses obtained in the questionnaires were indicative of what Personal and Social Development Indicators (IDPS by the Spanish acronym) are intended to measure. Thus, through an ethnographic study in 18 schools in different districts of the country, information was collected to enable the definition of the development levels for each of the IDPS and their respective dimensions. This allowed a comparison to be drawn and the level of consistency between the results reported by both methodologies to be defined, thanks to the provision by the Quality Education Agency of all the information related to the scores from the SIMCE questionnaire from all schools that took part in the case studies and the gross scores for the whole country.

This paper is divided into four sections. The first outlines the background information and conceptual elements considered for the research, mainly in relation to large-scale evaluation and the validity of instruments and the IDPS. Then, in the methodological framework, the general study methodology, the instruments used to collect information, and the aspects related to the specific methodology for consistency analysis are all explained, detailing how the levels of achievement were defined based on each of the methodologies and the process of analysis itself. The third section shows the results of the study regarding the consistency, demonstrating medium-high consistency between the results of the questionnaire and the ethnographic findings. Finally, general conclusions are presented for the study covered in the paper and regarding the use of consistency analysis as a method to perfect large-scale evaluation processes, in addition to the recommendations associated with them.

Background Information

Large-scale evaluation

Large-scale evaluation in Latin America stems from a moment in the region determined by the need to advance with the systematization of information from the efforts made to expand educational coverage and increase levels of schooling. This process, which took place in the 1980s and was already definitively institutionalized in the late 1990s, was intended to provide a technical correlation to the need to achieve efficacy in student learning at schools in a large number of Latin American countries.

Thus, toward the end of the 1990s it was possible to take "the first steps toward the development of basic technical capacities to design and implement large-scale evaluations and to undertake statistical analyzes of the data collected" (Arregui, 2008).

Nevertheless, as regards the use of this information and the need for improvement of the various evaluation systems, according to Arregui (2008), the focus has been on two aspects. On the one hand, there is the need to technically improve the instruments to collect the evaluations and their subsequent analysis and, on the other hand, and perhaps even more complicated, is how to communicate all this information to the different users of the system and the wider community (Arregui, 2008).

At the local level, these elements have marked the evolution of the assessment system since its beginnings in the early 1980s, since they have enabled the instruments and their meanings to be reviewed, as well as the form and methodologies used to analyze the information. In addition to the above, the need to provide information that is useful to educational establishments, that is, to enable them to take decisions to improve their institutional and educational management, has been highlighted constantly in recent years.

In 2009 the General Education Law (LGE by the Spanish acronym) was signed in Chile to replace the LOCE, which establishes the duty of the state to guarantee quality education and ensure that it is provided to all students residing in the country. The main objectives of the new law are to ensure that all students receive an education that gives them opportunities for their formation and integral development, to seek to ensure quality education for all, to ensure that all students achieve the general objectives and learning standards, and, finally, to implement a Quality Assurance System for Education.

The creation of this Quality Assurance System is a fundamental landmark in terms of understanding large-scale assessments, as solely academic aspects are no longer considered in the evaluation process because non-academic indicators are included that have a value in themselves and provide for a much more comprehensive look at the evaluation.

Within this structure, the Education Quality Agency has a central role in terms of contributing knowledge to the education system regarding the state of this new perspective of the concept of quality in light of its function to assess the degree of compliance with the Personal and Social Development Indicators.

Personal and social development indicators

In recent years, national efforts in terms of school education have included a strong focus on ensuring a quality education for all students. Within this framework, the Quality Agency is the body in charge of evaluating the achievements of students according to the degree of compliance with certain standards, which, in addition to the academic standards, include the IDPS. This is consistent with the General Education Law, which postulates that one of the objectives of education should be the development of attitudes, knowledge, and skills so that in the personal and social field students develop positive self-esteem, act according to the values and norms of civic coexistence, practice physical activities, and acquire habits of hygiene and personal care, among others. The IDPS are intimately related to the Objectives of Transversal Fundamentals (OFT), which define a series of knowledge, skills, and attitudes that are oriented toward the development of competences that are considered fundamental for personal development and to develop in the social, labor, and civic spheres (Castillo & Contreras, 2014).

These indicators respond to the inclusion of non-academic elements that are considered part of the child's development and which should be promoted through school education. Currently, social and emotional learning is considered a process in itself, through which students acquire and develop skills that support learning, positive behavior, and constructive social relationships (Kendziora, Weissberg, & Dusenbury, 2011).

At present, the need to develop the individual and social skills necessary for people to face the demands of a constantly changing reality is becoming ever more evident. For this reason, in addition to teaching literacy in letters and numbers, the school has acquired the role of teaching literacy in emotions, social skills, decision making, and the management of interpersonal relationships (Teruel, 2000). However, until now, the school has had an almost exclusive function to teach and transmit academic aspects, without having explicit responsibilities regarding the development of students' social skills or interpersonal and personal wellbeing. In this vein, Castillo & Contreras (2014) express the urgent need to redefine the role of educational establishments as agents that promote the development of subjective wellbeing, that is, they deliver tools to students for the construction of personal and social projects that make them feel satisfied. This need is heightened in our country because subjective wellbeing and the characteristics to propitiate it are another area in which inequality is expressed. The hypothesis behind this approach is that the more capacities a society offers for the development of subjective wellbeing, the greater the probability that its members will attain it. Educational establishments thus appear as a privileged place in which public policies can contribute to the formation of different capacities at the different stages of development of children and young people.

In regard to this, we have to consider that it is not the explicit or formal curriculum that determines students' personal and social development. The promotion of these capacities is more closely related to other aspects of the school environment, such as student-teacher interactions and the interactions between the students themselves, relationships that comprise informal education or the hidden curriculum, through which the teacher acts as an agent of socialization and as an educator of their students in a non-explicit manner, in relation to the methodology, educational styles, and the socio-emotional climate that is generated in the classroom (Álvarez & Bisquerra, 1999).

Specifically, the IDPS are defined as: "a set of indices that provide information related to the fulfillment of the General Education Objectives in the personal and social sphere, in a manner complementary to the results of the SIMCE test and the achievement of the Learning Standards" (Decree 381).

In this regard, measurement and reporting of the IDPS is aimed at advancing toward providing a comprehensive education. Thus, in accordance with Decree 381, the role of the indicators in the evaluation of education quality is:

- To provide educational establishments with relevant information regarding the personal and social environment of their students.

- To serve as an input for the classification of schools in the Performance Categories[1], which allow the identification of schools that require direction and guidelines for improvement.

- To contribute to the self-diagnosis conducted by schools and the educational improvement plans (PME) they implement.

- To contribute to the design and evaluation of public policies and thus provide relevant information for research.

Four of these indicators, which are evaluated based on questionnaires, are described below[2]. In this regard, the definitions shown below correspond to those established in Decree 381.

The first indicator is Academic Self-esteem and School Motivation, which, on the one hand, evaluates the self-perception and self-assessment of students in relation to their capacity to learn and, on the other, the perceptions and attitudes that the students have towards learning and academic achievement. This indicator is comprised by two dimensions, Self-perception and academic self-evaluation and School motivation.

The second indicator is the School Climate. This considers the perceptions and attitudes that students, teachers, and parents and guardians have regarding the existence of a respectful, organized, and safe environment at the school. This indicator is comprised by three dimensions, Respectful Environment, Organized Environment, and Safe Environment.

The third indicator is the Citizen Participation and Education indicator, which considers students' attitudes toward their establishment; the perceptions of students, and parents and guardians of the degree to which the school fosters the participation and commitment of members of the educational

---

1   The National Quality Assurance System, established in 2012 by law 20,529 (SAC Law), states that the schools will be classified into four categories according to an adjusted socioeconomic performance index that is calculated annually depending on the academic results of the schools and the results of the IDPS.

2   There are eight indicators defined by law: Academic Self-Esteem and School Motivation, School Climate, Citizen Participation and Education, Healthy Living Habits, Gender Quality, Attendance, School Retention, and the Rate of Qualification for the Professional Technical Modality in Secondary Education. However, the study on which this paper is based used the indicators measured with questionnaires

.

community; and the perceptions of students about the degree to which democratic life is promoted. The dimensions of sense of belonging, democratic life, and participation are considered.

Finally, the Healthy Living Habits indicator assesses the self-declared attitudes and behaviors of students in relation to healthy living and their perceptions of the degree to which the school promotes habits that are beneficial for health. The dimensions considered are Food Habits, Active Life, and Self-Care.

Methodological Framework

The study on which this research is based consisted of conducting an ethnography in 18 educational establishments in 11 districts of 4 regions of the country, where school dynamics were observed in relation to the Personal and Social Development Indicators (IDPS): Academic Self-esteem and School Motivation, School Climate, Citizen Participation and Education, and Healthy Living Habits. The time spent at each school was 25 days and the gathering of information included participant observation of various school spaces (schoolyard, classroom, dining areas, board meetings, teachers' meetings, and extracurricular activities, among others); group interviews with students, teachers, parents and guardians, managers; individual interviews with students, parents and guardians, teachers, and other key informants; and the review of institutional documents. In addition to this and specifically to supplement the information gathered for the indicator of Academic Self-esteem and School Motivation, the Test of Self-Esteem in the School-Student Context (TAE by the Spanish acronym) was applied to sixth grade students (protocol TEA-Alumnos). The implementation of this test is based primarily on the difficulty of identifying or assessing the self-esteem of students based on the qualitative activities previously defined in the framework of the research (group and individual interviews, and observations), because, unlike the other indicators, measuring self-esteem requires the use of individually-applied qualitative techniques (projective tests, clinical interviews, or pictorial techniques) and it requires significant amounts of time, both in terms of implementation and analysis, which is difficult in the school context and particularly within the framework of this study. Therefore, the possibility provided by the TAE was to apply a validated and standardized screening test for the level of teaching in the measurement of academic self-esteem (Marchant, Haeussler & Torreti, 2016).

This ethnography was generally aimed at describing and analyzing the school dynamics surrounding Self-Esteem and School Motivation, School Climate, Citizen Participation and Education, and Healthy Living Habits in sixth grade students[3]. In turn, with this information, we sought to assess the degree of consistency between the results of the ethnographic study and the results from the questionnaires on the context, perceptions, and attitudes regarding SIMCE, as stated in specific objective number 5 defined for the study: *to assess the degree of consistency between the dynamics that occur in the educational establishments and the results of the four OICs[4] in order to determine whether responses are indeed indicative*

---

[3]   The general objective of the study, defined in the tender bases, is the following: describe, analyze, and compare the school dynamics with the results obtained in the four educational quality indicators (OIC) measured using questionnaires of the context, perceptions, and attitudes in 6th grade, as a way of assessing the level of confirmation between the information governed by these instruments and the observed dynamics that exist in educational establishments.

[4]   At the date of the study, these indicators were known as Other Educational Quality Indicators (OIC by the Spanish acronym), which was changed to Personal and Social Development Indicators (IDPS).

of the aspects that the OICs are intended to measure. It is therefore this objective that was mainly addressed through the consistency analysis and it is directly related to what is presented in this article. In addition to this aim, five other specific objectives were also addressed in the study, which, although complementary to the consistency analysis, are not specified since they do not relate directly to the subject of this paper.

We used triangulation between qualitative and quantitative approaches (triangulation between methodologies) because it is a mechanism that enables the validity of the different methodologies used to be enhanced. The more methods and approaches used to study the same phenomenon, the greater the reliability and validity of the results. This prevents the results of a certain information-collection instrument from depending purely on the indicators used (Mathinson, 1998, Cohen, Manion, & Morrison, 2013), as well as enhancing the understanding of the phenomenon, assuming that the weaknesses of one approach are supplemented by the other (Valencia, 2013).

The fieldwork was carried out by a team of nine researchers, including the key team that led the study (i.e. project director, general coordinator, methodological adviser). The activities related to each of the 18 case studies were led by a principal investigator, who was responsible for collecting information and the case analysis. In addition, a second researcher (part of the key team) was involved in each case, fulfilling the role of assisting with certain activities, supporting the principal investigator regarding technical aspects, and validating and providing feedback on the case report. In order to incorporate feedback and validation of the preliminary findings from the school stakeholders, the results of each case were presented to the management team of each educational establishment. This also allowed the teams in each school to reflect on their current status and their management associated with the IDPS.

In this way, key strategies were addressed to enhance the validity and reliability of the study, these being: the use of various information sources (parents and guardians, students, education assistants, and teachers, among others); diversity of information-gathering tools to research the same topic; triangulation at the level of researchers (Álvarez, 2011; Valles, 2000); and, finally, opportunities to incorporate those involved in the key findings of the study by the feedback made to the management teams of each school.

In order to analyze the material produced in the schools or the body of the research, work was done based on content analysis, which was intended to organize the information collected based on codes that allowed various sources of data production to be gathered (interviews, observations, field notes, institutional documents, among others) based on common codes incorporating both latent and manifest elements. This process was carried out considering categories and subcategories defined a priori —in a matrix of dimensions, sub dimensions, and associated indicators— as part of the study. Based on this matrix, the information collection instruments were created, which were revised and received feedback on their formulation in a two-way process between the researchers and the data, considering emerging elements, organizing the data based on increasingly abstract and complex categories, establishing relationships and inferences, and maintaining constant dialogue between the fieldwork and analysis (Glaser & Strauss, 1999; Cáceres, 2003).

The consistency analysis is thus based on the analysis of cases and the integrated analysis emerging from that. For each educational establishment a narration was made that includes the various different perspectives and techniques used in the fieldwork, in order to provide a comprehensive account of each context with its particularities, seeking a well-based description and interpretation of the dynamics included in the four areas of the IDPS. The analysis was conducted —in an iterative manner— in line with the following steps: analytical reflection on the data and selection, organization and reduction of the data using as a focus the matrix of dimensions and key sub-dimensions under which each indicator is understood, where it reveals the aspects that have to be characterized.

In order to conduct the consistency analysis itself, we defined a methodology that enabled comparison of the quantitative results obtained by each school in the SIMCE questionnaires with the qualitative results of the ethnography. The methodology for the contrast analysis first involved defining the achievement levels (NL) for the four indicators and their respective dimensions, from both information sources (questionnaires and ethnography). The NL categories defined for both sources were: high, medium-high, medium-low and low. These achievement levels were defined using a different methodology for each information source, one qualitative and the other quantitative, as described in the sections related to levels of achievement.

Tools

As regards information gathering, specific guidelines were designed for each of the activities with the different stakeholders based on the general matrix of dimensions and sub-dimensions, which oriented both the collection of information and subsequent analysis. This matrix was based on an exhaustive documentary review and was validated with national experts in the different topics of the IDPS. Although the matrix considered the four indicators and their respective dimensions —as defined by the Ministry of Education— a dimension of context was incorporated for the three indicators. As already mentioned, in later stages, the matrix received feedback from the process of information gathering and analysis.

Based on this, the following tools were created:

–   Individual interview guideline for directors. Objectives: to observe where the school is in general terms and elements of its identity and history; to study the relevance that it gives to each of the IDPS indirectly; to problematize key issues in relation to the IDPS; and to see how they work, who works on them, and to what extent they are managed, neglected, or critical aspects on the part of the management.

–   Individual interview guideline for sixth grade teachers. Objective: to go deeper into the dynamics, practices, and perceptions regarding how the IDPS are realized, including important elements for comprehension and contextualization.

– Individual interview guideline for sixth grade students. Objective: to go deeper into the dynamics, practices, and perceptions regarding how the IDPS are realized, including important elements for comprehension and contextualization.

– Group interview guideline for sixth grade teachers, parents and guardians, and sixth grade students. Objectives: from the perspective of the different stakeholders, to go deeper into how they see the establishment and the major challenges; to explore perceptions and find out about the dynamics regarding the relevance and how the IDPS are realized at the school, identifying dynamics, specific practices, and incorporating critical vision in this regard.

– Guidelines for recording observations: for the classroom, teachers' room, school entrance and exit, and recesses. Objective: to identify specific dynamics and practices related to the IDPS, using as a guide the general IDPS observation matrix (record in field notes), along with the key indicators for each observation space.

In addition to the above, the Test of Self-Esteem in the School-Student Context (TAE) was applied to the sixth grade students. As already mentioned, justification for the application of this questionnaire in the framework of the study lay mainly in the understanding of self-esteem as the general assessment that people make of themselves, which is present in all aspects of life and mediates at all times between the stimuli that the person receives and the responses they make. Considering this, the results produced by the TAE were expected to be consistent with the results of the context questionnaire, on the understanding that the standardized test incorporates other components. One of the most important points that this standardized questionnaire reveals is the interaction of self-esteem with the context in which the dynamics develop.

Achievement levels based on the ethnographic study

The case studies carried out for each school enabled the identification of characteristics in the dynamics and perceptions related to the IDPS and their respective dimensions. Based on this information, an initial distribution was created for each of the 18 schools, classifying them by indicator and their dimensions in an achievement level (NL). That is, based on the qualitative analysis, each of the four indicators and their dimensions were associated, in each case, with a development level (high, medium-high, medium-low or low).

In order to assign an NL to each case from the fieldwork, we conducted an analysis that consisted of characterizing all the cases in the sample by case study, by indicator, and by dimension, summarizing the information provided by the researchers in a matrix of categories. Then two parallel processes were carried out for the analysis of each indicator. On the one hand, schools were classified from low to high in a certain indicator with respect to the general level of achievement, beginning by identifying the extremes that were characterized positively or negatively with respect to the rest of the cases. On the other hand, the qualities of each sub-dimension were analyzed between cases with similar locations in this classification, starting at the extremes. With this second process, patterns were defined, first between the positive and negative extremes, and then for the intermediate cases, which led to the identification of associated parameters. There was mutual dialogue between these two processes, that is

to say, the location of some cases was changed in the process of defining parameters, until agreement was reached and there was an explanation of the levels of achievement through the parameters. For this exercise, constant dialogue with the principal investigators was key, so as to contrast and validate the defined NL for each case, going beyond the information provided in the case report.

This allowed the analysts to form an overview of the dynamics that are shown for the indicator and its dimensions at the different schools. Thus, based on this process of analysis, a preliminary NL was assigned to each indicator and dimension for each of the 18 cases. As stated, this achievement level was then validated with the principal investigators who collected the information and created the case report for each educational establishment.

The above permitted the creation of a characterization of the school dynamics for each of the four levels of achievement, by indicator and their respective dimensions. Thus, from the general coding, levels of development were revealed for each Indicator of Personal and Social Development, which was made possible partly due to the size and heterogeneity of the sample.

Levels of achievement based on the results of the 2014 SIMCE questionnaire

For the information revealed by the Education Quality and Context Questionnaires, achievement levels were defined based on the percentiles identified for each case in relation to the universe of schools according to the following distribution: High NL, between the 70th percentile or greater and 100th percentile; NL medium-high, between the 50th percentile and 69th percentile; NL average low, between the 30th percentile and 49th percentile; and low NL, between the 0 percentile and 29th percentile (Table 1) For each educational establishment, a conversion was made from the gross score to the percentile —for indicators and dimensions— based on information provided by the Education Quality Agency.

Table 1
Achievement levels, questionnaires

High level of achievement (A)	Between 70th percentile and 100th percentile
Medium-high level of achievement (MA)	Between 50th percentile and 69th percentile
Medium-low level of achievement (MB)	Between 30th percentile and 49th percentile
Low level of achievement (B)	Between 0 percentile and 29th percentile

Methodology of analysis

The consistency in this study is understood as the level of convergence between the levels of achievement defined by both methodologies for an indicator and its dimensions. Considering this, the consistency is always obtained from the comparison between the information —specifically between the NL— produced by each of the sources, that is, by the ethnographic study and the 2014 Questionnaires on Education Quality and Context.

In order to distribute the cases according to the consistency, three levels were defined (Table 2): high consistency (+) when the NL defined for a case from the ethnography coincides with the NL defined for the same case from the results of the questionnaire; medium consistency (•) when the NLs defined from each information source do not coincide, but this divergence occurs between adjacent and low consistency achievement levels (-) when the divergence between the information provided by both sources of information is between NLs that are not adjacent (Table 2).

Considering all of the above, the consistency analysis uses the previous classification (high consistency, medium consistency, and low consistency), conducting a comparison of the achievement levels defined for each indicator between the methodologies. Specifically, we seek to answer the question about whether the results or NL assigned from the ethnography for an indicator coincide or differ from the NL defined for the same indicator based on the score of the questionnaires.

Table 2

Levels of consistency between ethnography and questionnaires

Level of consistency	Criteria
High consistency (+)	The achievement level defined for a case from the ethnography coincides with the achievement level defined for the same case based on the results of the questionnaires
Medium consistency (•)	The levels of achievement defined from each of the sources do not coincide, but this divergence occurs between adjacent achievement levels (Table 6)
Low consistency (-)	The divergence between the information provided by both sources occurs between achievement levels that are not adjacent (Table 6)

Discussion of Results

Table 3 shows the distribution of the sample in the different levels of consistency by indicator. In this table it can be seen that most of the cases show average consistency between the information provided by the ethnography and the results of the 2014 Questionnaires on Education Quality and Context. The percentage of schools that are situated in this intermediate level ranges between 56% and 67% of the total, with the School Climate indicator showing the highest percentage of cases with an average level of consistency. As defined in the methodology, this average level implies that, for a school, the achievement levels defined from each of the information sources differ, but this divergence occurs between adjacent levels (Table 2).

The high consistency group is the second most represented group, where the achievement levels coincide for the indicators reported by both sources. The composition of this group does not show a correlation as to which schools comprise the level for each indicator. Only one case showed a high consistency in three indicators, and four cases show this level of consistency in two of the indicators. Thus, of the 13 schools in the sample with high consistency in at least 1 indicator, only 5 have high consistency in more than one indicator.

The low consistency is the least represented group, with a percentage that varies between 7% and 19% (between 1 and 3 schools). Here only two establishments show low consistency for two

indicators, while the others show this level of consistency for only one indicator. The IDPS with the lowest percentage of cases at this level of consistency is the School Climate, so we may assume that the items of the tool related to its dimensions are more consistent with the ethnographic activities than in the case of the other three IDPS

The lower representation of cases with low consistency in the school climate indicator may be related to the incorporation of the context variable (for all indicators), where the level of importance achieved by the characteristics related to this new dimension varies depending on the indicator. Therefore, for the climate dimensions and variables, they are themselves contextual perceptions, so this fourth dimension does not add much value to the view from the questionnaires. This may partly explain the lower variability in this indicator, considering that the inclusion of the context dimension does not add new areas to investigate in relation to those that are addressed by the questionnaires. This analysis leads us to consider that the inclusion of the context variable, for each indicator, is relative to the approach already taken in the questionnaires.

Table 3
General level of consistency, by IDPS

Consistency level indicator	Academic self-esteem and school motivation		School climate		Citizen participation and education		Healthy living habits	
	N°	%	N°	%	N°	%	N°	%
High consistency	5	27%	4	26%	4	25%	6	33%
Medium consistency	10	56%	10	67%	9	56%	10	56%
Low consistency	3	17%	1	7%	3	19%	2	11%
Total	18	100%	*15	100%	16**	100%	18	100%

NB: *Case N° 3, Case N° 9 and Case N° 16 are not included in the analysis because they do not have results from the Personal and Social Development Questionnaires for the 6th grade course. **Case N° 7 and Case N° 16 are not included in the analysis, as they do not have the results of the Personal and Social Development Questionnaires for the 6th grade course

As shown in Table 3, specifically at the level of Academic Self-Esteem and School Motivation indicator, 56% of the cases show a medium level of consistency, that is, there is a distance between the achievement level assigned based on the ethnographic study and the achievement level produced by the questionnaires, but this is not highly significant. This implies that there are dynamics that do not coincide between the two sources of information, but there is a significant group of characteristics that do coincide. For the high consistency group (27%), the dynamics observed based on the ethnography show consistency with the results of the questionnaires, so they are classified at the same level of achievement based on both sources of information. For the three cases where we observe low consistency, we can also see that the results of the questionnaires show instability between their dimensions, that is, the difference between the scores of the two dimensions is greater than expected.

With respect to the IDPS on school climate, only one case in the total sample shows a low level of consistency, that is, the divergence between the levels of achievement from both sources of information is seen between levels that are not adjacent. The group with high consistency represents 27% of the

sample, while the average consistency is the one that is represented most, accounting for 67% of the cases. This allows us to conclude that, for most of the schools, the dynamics observed based on the ethnographic study differ to some degree from those associated with the quantitative results, but this inconsistency is not absolute. There is no correlation between the categories of consistency and the levels of achievement of the cases that comprise them.

What can be concluded for this indicator is that the achievement level assigned from the ethnographic study is always higher than that determined from the Personal and Social Development questionnaires, with the exception of the four cases with high consistency, where the levels of achievement coincide.

The results related to the Citizen Participation Education indicator show that, of all the cases, there are 25% that show a high consistency between the achievement level assigned from the ethnographic study and the achievement level defined from the results of the questionnaire. The largest group is represented by cases that have average consistency between their classifications by NL (56%), while the group with low consistency is comprised by three cases (19%).

Finally, regarding Healthy Living Habits, the highest percentage of cases are concentrated at the level of medium consistency (56%), which means that there is a difference between the achievement level defined from each source of information, but this is not an extreme difference. This means that although there are similar dynamics, it is possible to identify other characteristics that are not coincident. The group of schools with high consistency is comprised by six cases (33%) which are mainly in a low or medium-low NL for both sources of information. It should be noted that for this indicator the high category is represented to a lower degree in only two establishments based on the results of the questionnaires. For the two cases of low consistency, no correlation is observed regarding achievement levels, with the four levels being represented.

Hypothesis associated with inconsistency

There are certain general hypotheses that can partly explain the inconsistencies that are generated between the ethnographic observation and the results of the questionnaire. In first place, there is a purely methodological reason that responds to the nature of both sources of information. On the one hand, the questionnaires reveal elements that are specific and in accordance with the "needs" of the indicator, which is also part of an application subject to a certain time, without taking into account other contextual elements. The ethnography, on the other hand, implies a different process in terms of the time factor (observations were made during a certain period), the variety of information sources (observation, group and individual interviews), and a compression that necessarily assumes an important element of context.

This last point —the incorporation of a context variable in all the indicators based on the ethnography— is presented as a transversal explanation. Although the questionnaires address the issue of context, they do so in a somewhat tangential way and with a different intensity. In addition, the holistic and comprehensive view of the indicator may partly explain the differences compared with the results of the questionnaire. Finally, the self-reporting methodology of the questionnaires can lead to a subjective view of how the conditions of the context occur.

Another general hypothesis that is associated with the issue of inconsistency is the difference between stakeholders considered from each approach. For all its dimensions and indicators, the ethnography included the perceptions of teachers, parents and guardians, managers, and students, as well as the observation and review of secondary information. On the other hand, the questionnaires take into account more limited perspectives, which vary from one indicator to another. The inclusion of the perception of education assistants as key agents appears to be relevant. In the case of self-esteem and motivation, the questionnaire only includes the perception of the students. Here, the information supplied by teachers and parents and guardians enhanced and validated the observations and preliminary results that the researcher collected from classroom observations, recesses, and interviews with students.

Lastly, the comparisons between the levels of achievement assigned from both sources of information lead us to think that the weighting of the dimensions is not the same for each methodology. This implies that, based on the fieldwork, each dimension acquires a different prominence. In contrast, in the analysis of the questionnaires, the dimensions have fixed and equal weightings compared with one another.

Conclusions and Associated Recommendations

Considering the results presented above, we should note the fact that the general level of consistency between the results produced by the ethnography and the questionnaires was medium-high for both methodologies (between 56% and 67% of the cases). Based on this, the first conclusion indicates that, in spite of the results, there are aspects of the instrument that need to be refined further to achieve a higher level of consistency, with the aspiration of a high level for all four indicators.

Another conclusion that emerges from the study is the relevance of the theoretical definition of each indicator, and how the ethnography can feed back into it from the observation and analysis of the dynamics and perceptions, considering both sources to determine the questions contained in the questionnaires. Therefore, the methodology of the consistency analysis is a contribution in terms of explaining the need to consider the limitations associated with the conceptualization of the indicators, mainly based on the theory, as the possible causes of inconsistency. We should also consider the specific factor that was discussed in the results section regarding the weighting of each indicator in the final score that a school obtains in the questionnaires, a decision that may be guided based on what is observed in the ethnography. One of the central aspects regarding this point is the low consideration of contextual and management elements in the definition of each indicator when evaluating them in the school system.

The whole previous point is highly relevant to public policy, considering that behind the constructs there are biases in the conceptualizations. In light of this, the recommendation is to plan processes for the definition of indicators that include an exhaustive review of the literature, panels of experts including professionals with different technical perspectives, and qualitative studies that complement the definitions of the indicators, their dimensions and the weighting that will be given to each area. This is particularly relevant in the context of measurement of the IDPS, considering that the questionnaires

have so far been applied in a census-based manner and that they are indicators that are mostly in the process of being implemented in the school system, so the management teams are developing practices to handle them. This means that the comprehension of the indicators, and their scopes and dimensions (revealed in the questions included in the questionnaires) are highly relevant for comprehension and, therefore, for the management that the stakeholders in the school system conduct in relation to the IDPS.

Thus, regarding the scope and limits of the large-scale evaluation of non-academic indicators in the educational field, we can first draw a conclusion on the relevance of carrying out systematic investigations in which methodologies can be contrasted that give validity to the instruments used and allow them to be improved. Using the case of this study as a reference, the analysis of the field work allowed us to identify specific dimensions and indicators where there was inconsistency between what was observed in the ethnography and the results of the questionnaires, which produced recommendations for the modification, elimination, and incorporation of questions and dimensions, in addition to proposing changes among the stakeholders to whom certain items of the instrument apply. In relation to this, it is recommended that large-scale quantitative evaluations at school level be submitted to periodic contrast analysis in order to validate and enhance this methodology.

Finally, an extensive ethnographic study provides exhaustive information on how these indicators are developed in the school context, both in terms of dynamics and management. This is why the methodology fulfills a secondary objective of supplying information that should be considered as a basis for the development of standards and the definition of guidelines that allow the evaluating body to provide educational establishments with recommendations associated with their quantitative results by indicator and dimension.

References

Álvarez, M., & Bisquerra, M. (1999). *Manual de orientación y tutoría*. Barcelona, España: CISSPRAXIS, S.

Álvarez, C. (2001). El interés de la etnografía escolar en la investigación educativa. *Estudios pedagógicos*, *XXXVII (2)* 267 – 279.

Arregui P. (2008). Uso de los resultados de evaluaciones educativas a gran escala en América Latina. Grupo de Trabajo sobre Estándares y Evaluación del PREAL. Retrieved from http://www.colombiaaprende.edu.co/html/mediateca/1607/articles-174348_archivo2.pdf

Cáceres, P. (2003). Análisis cualitativo de contenido: una alternativa metodológica alcanzable. *Psicoperspectivas, II*, 53 – 82.

Castillo, J. & Contreras, D. (2014*). El papel de la educación en la formación del bienestar subjetivo para el desarrollo humano. Una revisión al caso chileno*. PNUD y UNICEF.

Cohen, L., Manion, L., & Morrison, K. (2013). *Research methods in education*. Routledge

Glaser, B., & Strauss, A. (1999). *The discovery of grounded theory: strategies for qualitative research*. New York: Aldine De Gruyer.

Kendziora, K., Weissberg, R., & Dusenbury, L. A. (2011). *Strategies for social and emotional learning: Preschool and elementary grade student learning standards and assessment.* Newton, MA: National Center for Mental Health Promotion and Youth Violence Prevention, Education Development Center, Inc.

Martínez, F. (2209). Evaluación formativa en aula y evaluación a gran escala: hacia un sistema más equilibrado. *Revista Electrónica de Investigación Educativa, (11)* 1-18. Retrieved from http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S160740412009000200002&lng=es&nrm=iso. ISSN 1607-4041.

Marchant T., Haeussler, I. M., Torretti, A. (2016). *Test de Autoestima del Escolar (TAE). Evaluación del alumno vía auto-reporte y vía profesor.* Fundación Educacional Arauco. Ediciones UC.

Mathison, S. (1988). Why triangulate? *Educational researcher, 17(2),* 13-17

Meckes L. (2007). Evaluación y Estándares: Logros y desafíos para incrementar el impacto en calidad educativa. *Revista Pensamiento Educativo, (38)* 351-371

Teruel, M. (2000). La inteligencia emocional en el currículo de la formación inicial de los maestros. *Revista Interuniversitaria de Formación del Profesorado, (38)* 141-153.

Valencia, M. M. A. (2013). La triangulación metodológica: sus principios, alcances y limitaciones. *Investigación y educación en enfermería, 18(1).*

Valles, M. (2000.) *Técnicas cualitativas en investigación. Reflexión metodológica y práctica profesional.* Madrid: Síntesis Sociología.