The harmonized joint pilot tests for 2010 Population Census Round in Latin American Countries: an integration experience

Alicia Bercovich
Instituto Brasileiro de Geografia e Estatística, IBGE
Rio de Janeiro, Brazil

1. Introduction

Since 2000 Census Round planning, six Latin American countries made a joint effort to harmonize a set of common variables, their concepts, definitions and classifications. The Census methodology was also partially harmonized, from the initial phases of Census work to results dissemination.

The Project goals were, among others, widening the cooperation between the countries in the production, use and dissemination of Census data and compiling socioeconomic statistics that would enable to make comparisons between the countries.

For the 2010 round, countries agreed on adding three new topics to be harmonized. The themes were selected based on users demand and relevance for all the participating countries: ***disabled population, international migration*** **and** ***indigenous people.***

Argentina, Brazil, Paraguay and Uruguay carried out several harmonized pilot tests on those topics as part of the 2010 Round Harmonized Population Censuses Project planning.

The disability and migration tests were carried out at the triple border between Argentina, Brazil and Paraguay on 2006 and 2007 respectively. Uruguay also carried out the test on 2008.

The test on indigenous population was carried out at indigenous communities belonging to the same ethnic group, both sides of Paraguay-Brazil border.

The field work was accompanied by non-participative observers representing NSO's of almost all South American countries and Mexico.

The paper addresses the development of the harmonization project since the 2000 census round, the Joint Pilot Test on Disability for the 2010 round as an example of the work done to harmonize concepts. There are also shown some results and an example of the comparative analysis done by the countries.

2. Background

The project for the economic integration of the Mercosur countries had shown the need to compile socioeconomic statistics that enable comparisons to be made between the countries.

The comparison of data had been difficult, mostly an aggregation of tables, considering the different contents and methodologies of the respective Censuses (IBGE, 1993; INDEC, 1995).

During the 90’s the NSO’s representatives met to evaluate the experiences of the last censuses, and the topics most frequently addressed related to the need for integration and comparability between the data produced by each country.

Representatives of the NSOs considered the possibility of harmonizing the Population Censuses from the planning stage, beginning the project of the Joint Population Census for the Mercosur countries, plus Bolivia and Chile[1].

The integration project began in May of 1997, during the initial phase of work on the censuses of the year 2000 round. From then on, a great effort was made to harmonize the variables, concepts, definitions and classifications, starting from the planning phase of the Censuses for the six countries. The project included the generation of a common micro data base with harmonized information on the six countries.

Since it would not be possible to harmonize all the variables of all the censuses, it was decided to select a set of common variables, and for those variables, to compatibilize definitions and concepts.

The strategic objective was, besides obtaining comparable statistical data, to optimize the methodological, human and technological resources of the countries for all the Censuses to be carried out in an harmonized way.

This also would allow obtaining Institutional support to be able to carry out the six Censuses as close to the year 2000 as possible, to make easier horizontal cooperation and to potentialize the technical international support

It was also considered important to try to achieve methodological homogenization, and to ensure comparability of the classification and coding systems. The goal was to prepare common tabulations and create a unique database.

To test the feasibility of the harmonization effort for the selected variables, Joint Pilot Tests were carried out at the borders between the countries, as described later.

2.1. Variables, Concepts and Classifications

Considering the heterogeneous nature of the census forms of each country, a subset of variables was selected to be harmonized for the Common Census:

- Some characteristics of dwellings, common to all forms, particularly those relating to sanitation conditions;
- The basic demographic characteristics of the individuals;
- Some educational characteristics, such as completed years of schooling and school attendance;
- The economically active and employed population, considering only the status in employment and Industry Classification; but after the enthusiastic work of the nomenclatures group, a preliminary common classification of occupations was also convened;
- Migration characteristics, especially migration between countries in the region.

2.2. Joint Pilot Tests

The first joint fieldwork experience was the first Joint Mercosur Pilot Test. This was initiated in November 1998 with the participation of Argentina, Bolivia and Brazil in border areas. The second took place between Argentina and Paraguay in 1999. The objectives of these experiments were to

[1] The members of Mercosur are Argentina, Brazil, Paraguay and Uruguay, but since the beginning Bolivia and Chile have been working with the group.

test the common content of the questionnaires of each country, test training methodologies, processing and construction of the common database.

The evaluation was done by the field teams of the participating countries and observers from the other countries. The repercussion of the tests surpassed expectations: the population of the regions included in the tests collaborated enthusiastically and the census activities were widely disseminated by the media.

2.3. **Achievements**

All the countries of the Project carried out their 2000 Round Censuses and considered the observer's comments for the design. The common classifications of Industry (CAES Mercosur) and Occupation were ratified by the countries.

One outstanding characteristic of the Project was the Horizontal Cooperation between the Institutes, through which experiences and technological progress were shared within a spirit of collaboration and great openness.

The attempt to achieve processing homogenization was partly successful: Uruguay, Argentina, Brazil, Bolivia and Chile used scanners for data capture and intelligent character recognition (ICR).

The countries established a technical cooperation network that included not only common methodologies but also the loan of scanners, support in capacity building, printing one country's test forms by another, mutual collaboration in automated/assisted coding, support between countries in scanning/ICR experience.

All the countries accompanied the experience of the Paraguay Indigenous Census and the census operations in Brazil, Bolivia, Argentina, Chile and Paraguay successively between 2000 and 2002.

It was agreed to create a Multidimensional Database, which is generated using the micro data from the censuses.

3. The Joint Pilot Test on Disability

The pilot test on disability was carried out at the border between Argentina, Paraguay and Brazil in 2006. The three countries tested the short set of questions proposed by the Washington Group for Disability Statistics. All the countries had carried out previously, cognitive testing of the original set of questions.

Non-participative observers representing NSO's of almost all South American countries and Mexico accompanied the fieldwork. Also experts on disability statistics from International Agencies were integrated to the observer's team. After the field work, the observers participated of an evaluation workshop to discuss their observations.

The three countries produced reports with analogous data analyses; this allowed comparing the results among them. Some of the analyses were discussed with the Washington Group, and proposed by its members as part of the harmonization task.

3.1. Results and Analysis

The analysis target was to test the consistency of the census questions proposed on disability, to evaluate how their interpretation may differ across diverse core domains, countries, and subpopulations.

This analysis complemented the cognitive testing already carried out by the three countries on early 2006 to gain deeper insight into how these core questions were understood by respondents.

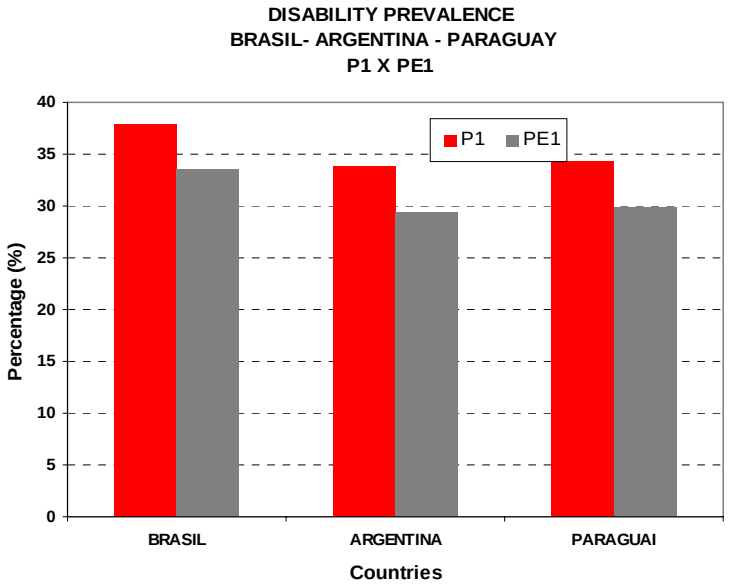
As the core questions proposed are very general, an extended set of questions could depict a more detailed picture of a person's level of functioning.

So, in order to get a better sense of what responses to the core questions indicated, the pilot test included some extended questions in the core domains in order to benchmark the responses to the more general questions.

The analysis consisted in comparing the answers to the short set of questions with the ones to the extended set, in a way that would allow to see if the core questions were good enough to identify people with disabilities in the investigated domains.

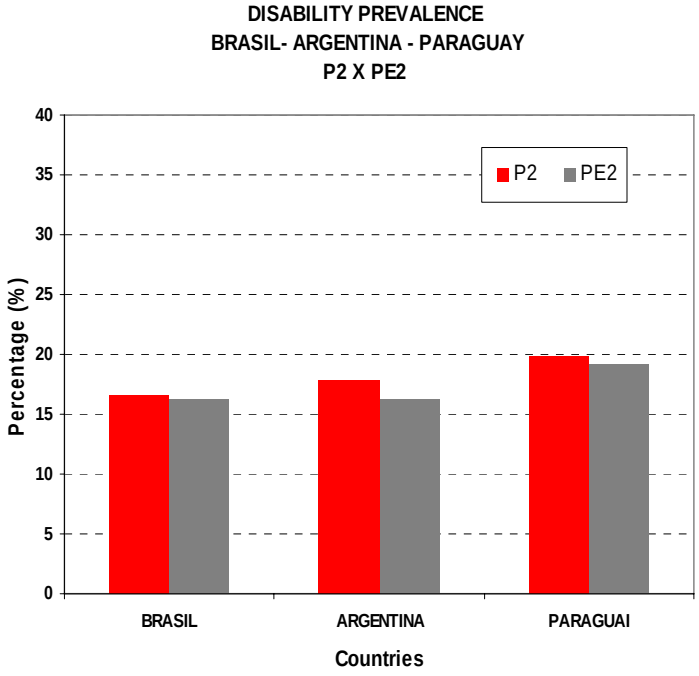
All the countries did the same analysis, which allowed comparing general prevalence of disability and also separated prevalence for each domain and severity level. Some of the results are shown below.

Figure 1



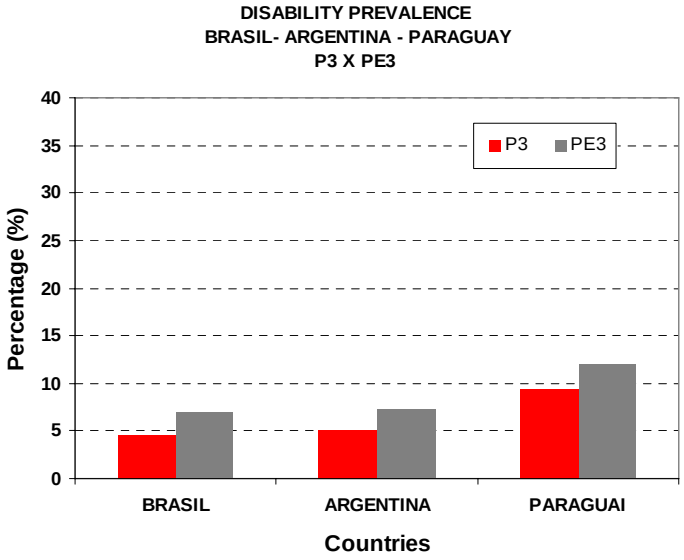
Source: I Joint Pilot Test on Disability: Brazil, Argentina and Paraguay, Nov. 2006.

Figure 2



Source: I Joint Pilot Test on Disability: Brazil, Argentina and Paraguay, Nov. 2006.

Figure 3



Source: I Joint Pilot Test on Disability: Brazil, Argentina and Paraguay, Nov. 2006.

The ***Figures 1, 2 and 3*** show the comparison between P (prevalence as measured by the core

questions) and PE (prevalence measured by the positive answers to the extended set of questions), for the three countries. The numbers P1, P2 or P3 identify level of severity: P1: if response was *some difficulty, a lot of difficulty, or can't do at all* in at least one domain; P2: if response was *a lot of difficulty or can't do at all* in at least one domain and P3: if response was *can't do at all* in at least one domain.

As can be seen, the core questions (red) were good to identify people with at least one disability (Figure 1), if compared with the people who answered positively to the extended set (grey). The proportion of people with at least one disability was bigger for Brazilians than for people of Paraguay or Argentina. For higher level of severity, Paraguay had higher proportion of persons in these categories.

The same analysis was performed by domain, and the core domains (vision, audition, mobility) had a good identification of disabled people trough the core questions. For the other domains, the best performance was for the questions on self care.

3.2. Correlation Analysis

To make easier statistical comparisons, the responses to the core questions were transformed from categorical variables to binary variables (Mont, 2005).

D1, D2, D3 were defined for each domain such that:

D1 = 1 if response is *some difficulty, a lot of difficulty,* or *can't do at all*, else =0
D2 = 1 if response is *a lot of difficulty* or *can't do at all*, else =0
D3 = 1 if response is *can't do at all*, else=0

As can be seen, D1 is the broadest definition of a disability and D3 is the most limited. The same was done for the extended set of questions.

The relationship between D1, D2, and D3 for the core questions and ED1, ED2, and ED3 for the extended questions was analyzed constructing the correlation matrix for each domain. Each cell is the correlation coefficient of the variables in the corresponding row and column. D0 and ED0 have been added, which represent people without reported difficulties in the core questions and in the extended set respectively.

Table 1

CORRELATION MATRIX BY COUNTRY AND DOMAIN
I JOINT PILOT TEST: BRAZIL - ARGENTINA – PARAGUAY

CORE QUESTIONS	SELF CARE											
	EXTENDED QUESTIONS											
	BRASIL				ARGENTINA				PARAGUAI			
	ED0	ED1	ED2	ED3	ED0	ED1	ED2	ED3	ED0	ED1	ED2	ED3
D0	**0,82**	-0,82	-0,75	-0,66	**0,61**	-0,61	-0,54	-0,49	**0,57**	-0,57	-0,56	-0,53
D1	-0,82	**0,82**	0,75	0,66	-0,61	**0,61**	0,54	0,49	-0,57	**0,57**	0,56	0,53
D2	-0,73	0,73	**0,83**	0,77	-0,64	0,64	**0,70**	0,64	-0,58	0,58	**0,63**	0,62
D3	-0,60	0,60	0,73	**0,85**	-0,60	0,60	0,74	**0,77**	-0,55	0,55	0,62	**0,66**

Source: I Joint Pilot Test on Disability: Brazil, Argentina and Paraguay, Nov. 2006.

Table 1 shows the correlation between the answers to the core questions and extended set for ***Self Care***, for the three countries and the three levels of severity.

The correlation coefficients in the matrix are used to test if there are statistically significant differences between the Di and EDi measures. If those differences were statistically significant then the measures are not equivalent. The comparisons between measures are done looking to the diagonal of the matrix for each one of the domains tested.

For Brazil, all the correlations Di vs. EDi are higher than 0,80 for *Self Care Domain*, and also Argentina and Paraguay had acceptable values, as it is shown in ***Table 1***.

Generally, the Domains of: Vision, Hearing, Mobility and Self Care, had acceptable levels of correlation for the three Countries. Communication and Cognition had lower correlations, corresponding to problems already pointed by the observers of the field work.

Table 2

CORRELATION MATRIX BY COUNTRY AND DOMAIN
I JOINT PILOT TEST: BRAZIL - ARGENTINA - PARAGUAY

CORE QUESTIONS	VISION											
	EXTENDED QUESTIONS											
	BRAZIL				ARGENTINA				PARAGUAY			
	ED0	ED1	ED2	ED3	ED0	ED1	ED2	ED3	ED0	ED1	ED2	ED3
D0	**0,65**	-0,65	-0,46	-0,17	**0,59**	-0,59	-0,45	-0,19	**0,29**	-0,29	-0,15	-0,03
D1	-0,65	**0,65**	0,46	0,17	-0,59	**0,59**	0,45	0,19	-0,29	**0,29**	0,15	0,03
D2	-0,45	0,45	**0,54**	0,23	-0,46	0,46	**0,60**	0,30	-0,14	0,14	**0,15**	0,05
D3	-0,01	0,01	0,03	**0,05**	-0,19	0,19	0,31	**0,74**	-0,03	0,03	0,05	**0,07**

Source: I Prova Piloto Conjunta sobre Deficiência, Brazil, Argentina and Paraguay, Nov. 2006.

Especially in Vision Domain, it can be observed, for Brazil, a very low correlation only between D3 and ED3, as shown in ***Table 2***. It was detected that the translation to Portuguese of "Cannot do at all" yielded to confusion with "No difficulty". The wording was modified for the next pilot test.

Other domains, like hearing, also presented the same problem of misinterpretation. The new wording was also introduced for those domains. The two pilot tests carried out afterwards with modified wording, showed that apparently, that problem was solved.

In the case of Paraguay, one of the questions on Vision was missing in the final version of the form, so the low correlations for Vision Domain could not be considered for evaluation.

Tables 3 and ***4*** show the Correlation Matrix for the Mobility and Cognition Domains respectively.

Table 3

CORRELATION MATRIX BY COUNTRY AND DOMAIN
I JOINT PILOT TEST: BRAZIL - ARGENTINA - PARAGUAY

CORE QUESTIONS	MOBILITY											
	EXTENDED QUESTIONS											
	BRAZIL				ARGENTINA				PARAGUAY			
	ED0	ED1	ED2	ED3	ED0	ED1	ED2	ED3	ED0	ED1	ED2	ED3
D0	**0,66**	-0,66	-0,57	-0,37	**0,78**	-0,78	-0,69	-0,45	**0,74**	-0,74	-0,74	-0,58
D1	-0,66	**0,66**	0,57	0,37	-0,78	**0,78**	0,69	0,45	-0,74	**0,74**	0,74	0,58
D2	-0,56	0,56	**0,67**	0,50	-0,66	0,66	**0,79**	0,59	-0,64	0,64	**0,77**	0,64
D3	-0,34	0,34	0,46	**0,65**	-0,37	0,37	0,48	**0,69**	-0,42	0,42	0,52	**0,64**

Source: I Prova Piloto Conjunta sobre Deficiência, Brazil, Argentina and Paraguay, Nov. 2006.

Table 4

CORRELATION MATRIX BY COUNTRY AND DOMAIN
I JOINT PILOT TEST: BRAZIL - ARGENTINA - PARAGUAY

CORE QUESTIONS	COGNITION											
	EXTENDED QUESTIONS											
	BRAZIL				ARGENTINA				PARAGUAY			
	ED0	ED1	ED2	ED3	ED0	ED1	ED2	ED3	ED0	ED1	ED2	ED3
D0	**0,46**	-0,46	-0,34	-0,22	**0,55**	-0,55	-0,42	-0,30	**0,52**	-0,52	-0,50	-0,40
D1	-0,46	**0,46**	0,34	0,22	-0,55	**0,55**	0,42	0,30	-0,52	**0,52**	0,50	0,40
D2	-0,47	0,47	**0,54**	0,39	-0,56	0,56	**0,63**	0,47	-0,49	0,49	**0,58**	0,50
D3	-0,29	0,29	0,46	**0,68**	-0,43	0,43	0,52	**0,70**	-0,36	0,36	0,48	**0,60**

Source: I Prova Piloto Conjunta sobre Deficiência, Brazil, Argentina and Paraguay, Nov. 2006.

4. Final Remarks

It was showed, in a simplified way, how some Latin American Countries are working together to carry out harmonized censuses for the 2010 Round.

The analysis of the Pilot and Cognitive tests on disability helped to identify the questions to propose to be included in the Census. The experience of joint tests in the border was again fruitful and allowed to discover similarities and differences among the countries, even when the same questions were applied in the field.

This harmonization strategy stimulated horizontal cooperation, solidarity between participants and an interchange of knowledge and experiences that had not been foreseen in the design of the original project.

This experience, which arose from the political will of the Institutes and the manifest interest of the technical staff in combining weaknesses to construct strengths, can be considered a paradigm of an efficient and participative working style. Part of the success is probably due to the spontaneity of the project, with the countries demonstrating openness to harmonize their efforts, no one taking the leadership.

BIBLIOGRAPHY

Bercovich, A. (1999). Integrating Concepts, Classifications and Knowledge: The experience of the 2000 Demographic Census for the Mercosur. Bulletin of the International Statistical Institute, Proceedings of the 52nd Session. Book 1. Helsinki.

Cazzolli, L. Bercovich, A. and Sá Vidal, R. (2004). Proyecto de Censo Común del MERCOSUR: una experiencia de integración regional. Proceedings of the United Nations Symposium on Population and Housing Censuses, United Nations Statistics Division, New York, September 13-14, 2004.

Colledge, M. J.(1997). Statistical integration through metadata management. Bulletin of the International Statistical Institute, Proceedings of the 51st Session. Book 1. 319-322.Istambul.

DGEEC, Dirección General de Estadística, Encuestas y Censos de Paraguay, 2007. Prueba Piloto Conjunta sobre Medición de la Discapacidad en los Censos de Argentina, Brasil y Paraguay de la Ronda del 2010, paper presented at the Workshop *"Ronda de Censos 2010. Una visión armonizada"*, CELADE, Santiago de Chile, June 2007

Giusti, A. (1998). El Censo del 2000 en el Mercosur: posibilidades de una empresa conjunta. Paper delivered at the session O Censo Demográfico do ano 2000: Perspectiva Nacional e Internacional. XI Encontro Nacional de Estudos Populacionais. ABEP, Cachambú, Brazil.

INDEC, Instituto Nacional de Estadística y Censos de Argentina , 2007. Análisis de los resultados de la Prueba Piloto Conjunta –Argentina, Brasil y Paraguay- para la Medición de la Discapacidad (PPC) Puerto Iguazú y Posadas (Argentina), Foz de Iguazú (Brasil) y Ciudad del Este (Paraguay), document presented at the Workshop *"Ronda de Censos 2010. Una visión armonizada"*, CELADE, Santiago de Chile, June 2007

IBGE (1993). MERCOSUL: Sinopse Estatística. Instituto Brasileiro de Geografia e Estatística. Rio de Janeiro.

IBGE, Comitê do Censo Demográfico 2000 (1998). Ata do Rio de Janeiro sobre o Censo 2000 no MERCOSUL, Dezembro 5 de 1997. In: II Seminário sobre o Censo 2000 no MERCOSUL. 9-13. Rio de Janeiro.

IBGE, Instituto Brasileiro de Geografia e Estatística, 2007. Report on the I Joint Pilot Test on Disability for the 2010 Demographic Census, I Prova Piloto Conjunta sobre Pessoas com Deficiência, paper presented at the Workshop *"Ronda de Censos 2010. Una visión armonizada"*, CELADE, Santiago de Chile, June 2007

INDEC Instituto Nacional de Estadística y Censos de Argentina (1995). MERCOSUR: Sinopsis Estadística. Instituto Nacional de Estadística y Censos. Buenos Aires.

INDEC, Instituto Nacional de Estadística y Censos de Argentina , 2007. Análisis de los resultados de la Prueba Piloto Conjunta –Argentina, Brasil y Paraguay- para la Medición de la Discapacidad (PPC) Puerto Iguazú y Posadas (Argentina), Foz de Iguazú (Brasil) y Ciudad del Este (Paraguay), document presented at the Workshop *"Ronda de Censos 2010. Una visión armonizada"*, CELADE, Santiago de Chile, June 2007

Mont, Daniel. 2005. Analysis Plan for Pilot Testing Census Questions, *The World Bank - UN Washington Group on Disability Statistics*, Washington, June 2005

Poulain, M. (1997). L'Harmonisation des concepts et des definitions comme etape indispensable de l'integration des statistiques. Bulletin of the International Statistical Institute. Proceedings of the 51st Session. Book 1. 315-318. Istambul.

United Nations (2008). Principles and Recommendations for Population and Housing Censuses (Rev. 2). New York.

Willeboordse, A. (1997). Discussion paper of the IP 27.- Integration of Statistical collections to improve consistency. Bulletin of the International Statistical Institute. Proceedings of the 51st Session. Book 3, 162-163. Istambul.