http://dx.doi.org/10.1590/2238-38752023v1329

I Universidad Nacional de La Plata, La Plata, Argentina
juan.piovani@presi.unlp.edu.ar
https://orcid.org/0000-0002-8774-3258

II Universidad Nacional de Misiones, Posadas, Argentina
baranger.denis@gmail.com
https://orcid.org/0000-0002-2501-4949

Juan Ignacio Piovani[I]
Denis Baranger[II]

COMPARABILITY AND EQUIVALENCE IN SOCIAL RESEARCH

INTRODUCTION

This study discusses some problematic issues around comparability and equivalence in empirical research, taking as a reference a cross-national study on the linguistic skills and the use of foreign languages, especially in the scientific production of researchers from Argentina, Brazil, and Chile[1].

To carry out this study, a team of scholars from these three countries worked collaboratively to design and implement a survey[2] based on an online self-administered questionnaire. The questionnaire was composed of several modules with queries about basic sociodemographic characteristics and educational trajectories; academic background and institutional affiliation, and various aspects linked to the acquisition of proficiency in foreign languages — mainly English — and their use for academic purposes: comprehension of foreign scientific literature, international mobility and training, participation in scientific events, publication in specialized journals, among others.

Even though the questionnaire applied in each country was basically the same, it was necessary to make some minor adaptations to account for local specificities, which later affected the possibilities for comparing its results. In turn, the definition of the target population and the "volunteer samples" resulting from the online self-administration of the questionnaire also posed challenges regarding comparability and equivalence.

This study takes up these matters in greater depth and, based on examples drawn from a cross-national study, illustrates some important conceptual and operative aspects related to comparability and equivalence in empirical research. Although this grounded discussion constitutes the central compo-

nent of this study (to which it devotes greater attention), it also includes two brief introductory sections meant to frame its examples in broader epistemological and methodological debates.

Following a logical sequence, the first section describes the status of comparison in the social sciences — and in science in general —, whereas the second one addresses problems of comparability and equivalence, especially in the case of cross-national studies. Finally, the third section discusses these issues in the light of concrete empirical examples arising from research into the linguistic skills of Argentinian, Brazilian, and Chilean academics and scientists.

COMPARISON IN SOCIAL RESEARCH

The status of comparison in science, and in social sciences in particular, has been the subject of intense epistemological and methodological debates. Especially in Political Science, an important intellectual tradition has put forward the idea of comparison as a scientific method: the comparative method (Archenti & Piovani, 2018; Piovani, 2001).

Both the expression and the idea of a "comparative method" have been widely disseminated since the 1960s. One of the clearest formulations of this perspective is due to the Dutch political scientist Arend Lijphart (1971; 1975). In his very influential 1971 article, *Comparative Politics and the Comparative Method*, Lijphart defends the status of a method for comparison, contesting other classical authors such as Lasswell (1968), for whom the idea of an independent comparative method in the field of political studies was redundant; Eisenstadt (1968), who argued that comparisons, rather than referring to a method in their own right, configure a particular focus within social analysis; and Almond (1966), for whom the comparative method is equivalent to "the" scientific method, rather than being just one within a variety of methods.

Lijphart's starting point is the assumption that the main objective of science is to discover general empirical relationships among variables. To achieve this goal, he acknowledges four different methods: experimental, statistical, comparative ones and case studies. This classification involves a hierarchical organization according to the degree of relative effectiveness of each method to verify/falsify hypotheses (Fideli, 1998), in which the comparative method occupies the third place, after the experimental and statistical methods, respectively, and before case studies. For this author, the logic of the comparative method is the same as that of the experimental and statistical methods, and it actually resembles the latter "in all respects except one": the comparative method deals with a *small number* of cases (Lijphart, 1971: 684) and, usually, *many* variables.

The classification of methods proposed by Lijphart has been the object of strong criticism, particularly as it is considered too restrictive to assign

science a single function, defined as establishing general empirical relations between variables. But even if we were to accept this epistemological stance, the idea of a "comparative method" itself has also been contested since, regardless of the method, comparisons play a fundamental role in any procedure used for determining the relations between variables.

In fact, in order to empirically assert these relations, it is necessary to "measure"[3] the state of the object under study in the variables at stake, and this cannot be achieved without comparisons, no matter whether this measurement occurs within the framework of a procedure that later enables the identification of the causal direction of these relations — such as an experiment — or whether it is limited to the analysis of correlations or associations by statistical techniques. Strictly speaking, even the simplest univariable description presupposes comparisons between objects in a variable that may show different states, values or modalities. In this sense, the mere assertion of a property implies the possibility of comparisons.

In short, many authors have argued that comparisons are inevitable in all scientific methods. Durkheim, in *Rules of the Sociological Method*, pointed out that comparative sociology was sociology itself, rather than one of its branches. Marradi (1991) has stated that it makes little sense to describe comparisons as a particular method of science. According to Ragin (2014), virtually all empirical social research involves some kind of comparison, and Swanson[4] is even more categorical when he claims that thinking without comparing is unthinkable.

However, even if its presence is in one way or another unavoidable, comparisons have varying places and significances across types of investigation. In fact, we could identify particular studies whose key cognitive objective implies an explicit and conscious comparison, in Sartori's[5] words. In these cases, both the design and practice of research are essentially defined, in all their dimensions and phases, by their central objective of comparing, thus requiring a theoretical framework that gives meaning and direction to the systematic comparison of certain units on determined properties, as well as a series of specific methodological decisions and technical instruments that enable the operationalization of such a comparison. Therefore, rather than configuring a method in its strict sense — what is known as the comparative method —, it constitutes a type of investigation that could be categorized as "comparative studies" or "comparative research."

Along these same lines, Fideli (1998: 12) states that "criticizing Lijphart's proposal does not mean denying the relevance of research practices that are usually designated with the label of comparative method," particularly when it refers to some specific styles of social research in which complex structures or systems — linguistic, cultural, institutional, social, political, educational, etc. — are confronted on the basis of their states in global properties (for example: the degree of structural differentiation, the degree of political stability, etc.)

or using data referring to different societies, cultures, nations, or even to sub-national units.

Although there is a tendency to almost automatically consider all comparative research to be international, the most widespread methodological definitions in no way exclude intra-societal comparisons, especially if it they can be based both in spatial and/or temporal dimensions. In this sense, Hantrais (2009: 2) defines comparative research as the study of "societies, countries, cultures, systems, institutions, social structures and change over time and space, when they are carried out with the intention of using the same research tools to compare systematically the manifestations of phenomena in more than one temporal or spatial sociocultural setting."

This same definition enables the acknowledgement that comparative studies go beyond a single discipline or methodological approach. Indeed, the broad field of social sciences and humanities have well-established traditions of comparative studies (both quantitative and qualitative) in disciplines such as Political Science, Sociology, Economics, History, Anthropology, and Linguistics, among others (Apter, 1971; Hantrais, 2009).

Beyond the multifaceted discussions around the comparative method and comparative studies, to define specific types of comparative research, the criterion of "dissimilarity of contexts" has been often used, giving rise to what Fideli (1998) calls "trans-contextual comparison." In turn, this encompasses two of the most classic and prevalent types of comparative research in the social sciences: cross-national and cross-cultural studies, which imply, respectively, the systematic comparison of nations and cultures[6].

International comparative studies of the cross-national type can be traced back to the mid-17th century — long before the expression 'cross--national' was even coined — to Hermann Conring's *Staatenkunde* and, along the same lines, to the *Statistik* School Achenwall founded at the University of Göttingen in the 18th century (Piovani, 2006). Revitalized after the Second World War under the framework of Comparative Politics, and in particular thanks to influential studies such as Almond and Verba's *The Civic Culture*, they became even more important from the end of the 1980s onward, as Hantrais (2009) maintains, due to the growing interest in international projects and networks.

Although the use of the term 'cross-national' has been standard in Political Science to refer to this type of study, international comparisons in other disciplines have frequently avoided this label on the assumption that the prefix 'cross' implies a compromise with the functionalist idea that the objects under study are functionally equivalent in different contexts and on the supposed restricted nature of these investigations, which tend to be limited to the mere juxtaposition of data (Hantrais, 2009). Yet acknowledging the formidable difficulties of achieving rigor in cross-national surveys, Jowell (1998) claims that their scientific standards are often poor and that they tend

to propose heroic conclusions based on faulty data, lacking the sort of explanations and interpretations that would enable them to go beyond the simple description of differences between nations.

However, as Hantrais (2009: 2) points out, "although this assessment may have been justified with reference to the macro level research carried out in the 1950s and 1960s, it is manifestly not necessarily the case in many of the cross-national comparative studies undertaken in later years." In any case, the expression 'cross-national' remains largely pervasive in international comparative research, totaling over 2 million occurrences in Google Scholar as of June 2020.

Although cross-national research goes beyond a particular methodological approach, in general, secondary analyses of statistical data has tended to be rather dominant, or at least more widespread (probably because international organizations extensively promote them). Nonetheless, international statistical comparisons can also use primary data within the framework of studies that include the design of survey instruments (samples and questionnaires) as well as the administration of questionnaires and the systematic analysis of the resulting data, as is the case of the research on language skills we discuss in this study to illustrate the typical problems of comparability and equivalence.

THE PROBLEMS OF COMPARABILITY AND EQUIVALENCE

From a methodological point of view, the central problems in comparative research (especially of the trans-contextual type) arguably refer to comparability and equivalence. The latter becomes apparent if we understand comparisons as the intellectual operation of collating the states of one or more objects in one or more properties and in one or more points of time[7] (Archenti & Piovani, 2018, Marradi, 1991). Osgood captured this problem in a simple and straightforward way: "when is the same really the same? When is the same really different? When is the different really different?"[8]. This issue affects decisions and practices such as the identification and characterization of the objects to be compared, the definition of populations and units of analysis, the selection of universes and samples, the use of concepts and terms, the selection and "measurement" of indicators, and the collection and analysis of the data.

Despite the currently profuse literature on the subject (Ariely & Davidov, 2012; Davidov et al., 2014; Mullen, 1995; van Deth, 1998, 2009), recognition of the importance of these issues actually accompanied the historical development of cross-national research. The classical comparative studies of the 1950s and 1960s have received ample criticism for largely neglecting problems of comparability and equivalence and for having taken the social and political processes of the developed north as a universal parameter of comparison.

But as the risks of eurocentrism and ethnocentrism in this field of study became increasingly recognized, comparability and equivalence — as well as the specific means of ensuring them — commenced to be extensively addressed in the branch of social science methodology specialized in cross-national and cross-cultural studies. Indeed, as Miller, Slomczynski, and Schoenberg (1981: 178) claim, "the substantive equivalence of survey instruments used in comparative studies is receiving increasing attention."

Nevertheless, as far as comparability is concerned, the debate has often been polarized (George, 1986). Philosophers such as Winch and Quine have rejected the possibility of cross-cultural comparison. Winch (1964: 307) has challenged the likelihood of intercultural translatability on the basis of our inability "to share and engage in practices which it is peculiarly difficult for us to comprehend," whereas from a linguistic perspective, Quine's indeterminacy of translation thesis underlined the impossibility of completely translating the meaning of a word from one language into another. On the contrary, other authors have defended the feasibility of conducting scientific comparative research. According to George (1986: 167), Deutsch's improved accuracy of measurement, Levison's testing of the equivalence of meaning, and MacIntyre's proposal to move to an "appropriate and lower level of generalization than scientific laws" exemplify how the scientific method could be used (although conditionally) to conduct comparative research.

In any case, "guaranteeing that the measurement of a relevant construct is equivalent across contexts [...], constitutes a central concern when applying theories and instruments across different countries and/or over time" (Ariely & Davidov, 2012). But even before assessing the equivalence of measurements and determining cross-national differences and similarities based on them, "the investigator must establish that the measured variables in each country are sufficiently similar in content to be categorized as the same phenomenon and warrant comparison at all" (Miller & Slomczynski & Schoenberg, 1981: 178). Therefore, regarding equivalence, we face two closely related but different problems: the equivalence of phenomena under study in varying contexts and that of the instruments researchers develop to measure such phenomena, as well as resulting measurements. To a large extent, these issues relate to classic problems of validity and reliability. However, if validity is understood as a property of the relation between an indicator and the construct it is meant to "indicate," in the case of cross-national and cross-cultural studies, it is not just about determining such relation in one particular social and cultural setting, but in various settings in which the "same" indicator might "indicate" different things. Paradoxically, "political, linguistic and cultural differences that are fundamental for comparative analyses are a source of threat to the validity of measurements in any comparative analysis" (Ariely & Davidov, 2012). As far as reliability is concerned, some of its main challenges derive

from the fact that individuals in different contexts may understand the same question in vastly different ways (Brady, 1985).

According to Blasius and Thiessen (2006: 229), "much is already known about problems, such as response sets and response styles, unequal expertise by the data-gathering institutions, and heterogeneous interpretation of the meaning of a given question or of the response alternatives", but much less is known "about how to assess construct equivalence between the countries." Indeed, the methodological literature has thoroughly dealt with issues concerning survey questionnaire design, translation, and linguistic equivalence (Deutscher, 1968; Converse & Presser, 1986; King et al., 2003; Davidov & De Beuckelaer, 2010; Harkness et al., 2010). But as Mullen (1995) has argued, measurement equivalence shows three aspects — translation, metric, and calibration equivalence — that are necessary to establish cross-national reliability and validity. More recently, we have witnessed the development of sophisticated multivariate statistical techniques to diagnose cross-national equivalence (Mullen, 1995; Billiet, 2003; Ariely & Davidov, 2012). But beyond the indisputable potential of these and other techniques to improve the quality of comparisons in empirical research, great caution must be taken to avoid confusing a substantive problem, such as equivalence, with the technical mechanisms that seek to determine and measure it. In this sense, it is worth considering an analogy with the assessment of reliability: all too often, and due to the characteristics of the most used techniques for this (split-half, test-retest, etc.), reliability is regarded as a problem of internal consistency, obscuring the fact that it actually is a property of the relation between an operational definition and the empirical result it produces when applied to a specific case. Needless to say, internal consistency could be perfectly achieved even in cases in which the operational definition fails to produce a value that reflects the "true" state of a given unit of observation in relation to a determined indicator.

Moreover, the problem of equivalence of indicators may be conceived in other ways. So did Bourdieu, holding a completely original view on this topic, which we should consider at some length[9]. His argument was grounded on structuralism: as language was a system of differences, as Saussure conceived it, in society, individuals and groups distinguish themselves by properties whose relations constitute the social space. Although Bourdieu seemed to have foremost settled this idea in relation to his conception of social classes, it can be extended to other indicators[10].

In the first pages of *Distinction*, Bourdieu stresses the necessity to overcome the "positivistic arbitration of the so-called facts" as far as those are conceived as simple bi-variable relationships. Indeed, "(h)idden behind the statistical relationships between educational capital or social origin and this or that type of knowledge or way of applying it, there are relationships between groups maintaining different, and even antagonistic, relations to

culture" (Bourdieu, 1984: 12). This is the basis for a criticism of Lazarsfeld and his abstract empiricism, a critic which is grounded both in Bourdieu's theory of action and in his conception of the social space.

In each particular statistical relationship, we should determine exactly what its terms designate, otherwise it will remain devoid of meaning. To concentrate efforts on refining the measurement of the 'intensity' of the correlation is useless insofar as it involves accepting "the *illusion of the constancy* of the variables or factors resulting from the *nominal identity* of the 'indicators' (whatever they may indicate) or of the terms which designate them" (Bourdieu, 1984: 18). For each statistical relationship in fact contains a whole system of variables which determines its intensity.

The indicator should not be taken for its intrinsic value but for what it is representing, which of course, is related to the very nature of what is meant by an 'indicator.' The crucial point is that what the indicator 'indicates' (what it represents), can differ as the case may be for each correlation and in each context. Otherwise, we shall falling into that "positivistic faith in the nominal identity of the indicators" (1984:22).

It is not only that the same *name* may be concealing different realities — as it is frequently the case, in comparisons from one to another country (or from an epoch to another), in which concepts such as 'unemployed' or 'poor' have been the object of different operational definitions for statistical purposes —, but also that all operational definitions face a far more basic problem: the fact that indicators are unable to operate in isolation from others. Each indicator 'indicates' from its position in a system of indicators.

The originality of Bourdieu's position stands in his structuralist view of the problem: each indicator should be considered in the signification it acquires relationally as part of a system. The meaning of an indicator, far from being intrinsic (or apt to be established once and for all) depends entirely on its position in the system of differences that determines all possible indicators in the same way that Saussure conceived language as a system of differences in which the linguistic value of a sign depended on the relations it had with other signs. Bourdieu extends the structuralist argument about the differential value (in a linguistic sense) of signs to a methodological level. Hence, this contextual determination of the value of each indicator is what makes a prior analysis of their social significance essential so the results of a survey can be properly sociologically read.

Bourdieu identifies standard statistical analysis as a form of substantialism since "it treats the properties attached to an agent – occupation, age, sex, qualifications – as *forces* independent of the relationship within which they act. This eliminates the question of what is determinant in the determinant variable and what is determined in the determined variable, in other words, the question of what, among the properties chosen, consciously or

unconsciously, through the indicators under consideration, constitutes the *pertinent property* that is really capable of determining the relationship within which it is determined" (1984: 22). However, as Pinto (1999: 139) stated: "Only for methodological convenience, and provisionally, it can be said that diplomas determine practices. Qualifications do not work, nor do profession, income or wealth"[12].

Lazarsfeld's "variable analysis" can thus be seen as an improper extension of experimental reasoning to the historical world, in which manipulation tending to isolate the effect of each variable separately results in volatilizing the structure likely to give account for these effects. If the social world is always described as variable configurations, it is with these configurations we shall have to work. To the independent variable must be opposed a *system of variables*[13].

Standard analysis can be useful as an approximation to the object but, taken to its ultimate consequences, ends up producing the ancient confusion between "the things of logic and the logic of things." Thus, "(t)he particular relations between a dependent variable (such as political opinion) and so-called independent variables such as sex, age and religion or even educational level, income and occupation tend to mask the complete system of relationships which *constitutes* the true principle of the specific strength and form of the effects registered in any particular correlation. The most independent of 'independent' variables conceals a whole network of statistical relations which are present, implicitly, in its relationship with any given opinion or practice" (1984:103). These relationships ultimately constitute the social reality. Faced with the nominalism of the variables that neither really act nor create the world, Bourdieu opted for a strong version of ontological realism.

Taking all these preceding considerations as our background, the next section will return to some of the issues related to comparability and equivalence in cross-national research, illustrating them with concrete examples arising from an investigation on scientists' linguistic skills. More specifically, we will discuss some of the challenges tackled in this research, the strategies used to address them, and their limitations.

COMPARABILITY AND EQUIVALENCE IN A STUDY OF SCIENTISTS' LINGUISTIC SKILLS IN ARGENTINA, BRAZIL, AND CHILE

As mentioned in the introduction of this study (and in the last paragraph of the previous section), the main objective of this study (and of this section in particular) is to discuss some key issues of comparability and equivalence in cross-national empirical studies — especially regarding samples and indicators —, taking examples from research on linguistic competences of scientists from Argentina, Brazil, and Chile.

Therefore, this section, rather than describing the "contents" or "results" of these comparisons between the three countries, discusses some methodological problems that had to be addressed to make comparisons feasible or to interpret them beyond the face value of empirical indicators, which could have led to a misleading or biased analysis.

Almeida et al. (2022) thoroughly describes and discusses the results of these comparisons with a detailed analysis of the involved variables. The study deals with language proficiency acquisition, the relation between linguistic skills and scientific production, and the role disciplinary differences played in the use of foreign languages in the three mentioned nations.

One of the first problems this research faced was the comparability of scientist populations since the three countries have different national science and technology (Sci-Tech) systems. This problem originates in the very conceptualization of "scientist," which, in principle, might seem somewhat self-evident. However, each country (and comparisons among them) understands "scientist" differently, giving rise to some vagueness in their definitions.

While Argentina has a specific professional career titled "scientific researcher" within its National Scientific and Technical Research Council (CONICET) — a body independent of universities and linked to a national network of research institutes and centers —researchers in Brazil and Chile work mainly at universities and thus belong to their academic staff. In any case, they receive *ad hoc* funds from national Sci-Tech organizations such as the National Commission of Scientific and Technological Research (CONICYT)[14] and the National Council for Scientific and Technological Development (CNPq) to conduct research projects approved under a competitive basis.

To increase comparability, we decided to work with "matched" populations. The matching criterion consisted of meeting a series of approximately equivalent conditions regarding the definition of "scientist." The intention was to study the "most qualified" components of the scientific system and, consequently, those most prone to internationalization. However, the choice of the target population in each country was carried out pragmatically, resulting in groupings whose equivalence is far from perfect. In the case of Argentina, the target population was defined by all CONICET researchers; in Brazil, by the whole academic staff affiliated to "high-quality" doctoral programs (those with the best assessment results, as defined by the Coordination for the Improvement of Higher Education Personnel — CAPES); and, in Chile, by all researchers with projects funded by the National Scientific and Technological Development Fund (FONDECYT) — the most prestigious and competitive within CONICYT — from 2000 to 2014.

Given the limited size of the target populations, and considering the high non-response rates that characterize self-administered online surveys, it was initially decided, rather than resorting to a sample, to send the

questionnaire to all population units. As expected, only a small proportion of potential respondents completed the questionnaire, leading to the configuration of a (non-probability) volunteer or self-selected sample (strictly speaking, three samples, one from each country). Out of the total target population, 23.8% of Argentinian scientists responded the survey, compared to 8.5% in Brazil and 3.8% in Chile.

Moreover, the distribution of key variables within each sample differed across countries. For example, the Argentinian sample, which is to a large extent representative of the distribution of the population across variables such as age, gender, and scientific discipline, has a much greater proportion of young and female scientists. Therefore, apart from differences in defining target populations, the effective samples resulting from responses to the online self-administered questionnaire added likely biases to cross-national comparisons.

The methodological literature in social sciences has extensively analyzed the weaknesses of online surveys samples. The need to turn to statistical adjustments to improve representativeness and compensate for the non-random selection of respondents (such as post-stratification weighting or propensity score matching) is widely accepted (Brick, 2011; Lee & Valliant, 2009; Heen et al., 2014; Schonlau et al., 2007), even though the effectiveness of these adjustments has also been contested (Loosveldt & Sonck, 2008).

In the case of this investigation, sociodemographic variables (and quite possibly other variables, such as scientists' discipline or country of completion of their doctorate), could have been useful to construct matched subsamples, enabling comparisons controlled by an important series of intervening factors. However, the size of the resulting sample for each country — particularly regarding Brazil and Chile — was too limited to construct these matched subsamples. Thus, comparative analysis had to consider the aforementioned differences of gender, age and other relevant variables as likely explanatory factors. That is to say that some of the comparative differences observed regarding linguistic competences could well be, at least in part, the consequence of the different distributions of key variables in our self-selected samples rather than a "real" variation in the scientific populations of the three countries.

Beyond the pertinent discussion about the possibility of generating valid knowledge from samples such as those obtained in this study, the obvious conclusion is that the nature and quality of samples affect comparability. It is worth noting that this problem goes beyond non-probability samples since factors such as sample size (especially that of sub-samples of social groups subject to comparison), sample error, and response rate could also undermine the comparability of probability samples.

On the other hand, achieving comparable indicators in all three countries has not always been an easy task. As noted in the previous section, comparative

surveys often face difficulties to operationally define their indicators to ensure uniform understanding in varying settings. Despite the relative similarity of national contexts — given the geographical proximity and the (by some means) shared historical experiences of Argentina, Chile, and Brazil —, our research also confronted problems of translating/adapting its questions and, furthermore, the format of participants' responses. In many cases, the equivalence of categories proved especially complex.

The basic principle that presides over any possibility of comparison is that all subjects must undergo identical *stimuli* (questions or items). An aspect logically associated with it but usually overlooked (perhaps because it is much more difficult to ensure) refers to the aforementioned need for subjects to decode all questions or items the same way[15]. Even a slight change in their formulation may suffice to render answers no longer strictly comparable (Converse & Presser, 1986).

For example, in the study under consideration, the wording of the question regarding the moment of acquisition of foreign languages was not exactly the same: in Chile, participants were asked about the moment in which acquisition started; in Argentina, about those in which such skills were developed. The conceptualization underlying these questions is different since the Argentinian version assumes that the acquisition of a foreign language is a process that evolves over time. Therefore, respondents were able to select many categories (multiple response). The Chilean version, instead, concerned itself with the precise moment in which the learning of the foreign language begun, thus allowing for just a single reply.

The use of non-homogeneous criteria to formulate a given question and/or to define response categories in different versions of a questionnaire employed in a cross-national study evidently affects the possibility of establishing direct comparisons based on the face value of its answers. In other words, it is impossible to uncritically compare the percentages of both samples in categories that are only "apparently" equivalent. In the case of the evaluated study on linguistic skills, and due to the characteristics of the operational definitions used in Argentina and Chile, the higher percentage of Argentinians who reported having learned English during childhood and adolescence is unable to be matched with the lower percentage of Chileans who claimed the same and from this to linearly conclude that the acquisition of English occurs earlier in Argentina.

Another problem arose when analyzing researchers' social origin. Due to the inability of using the occupation of respondent's parents, it seemed reasonable — as a proxy — to resort to the latter's educational attainment to develop an index of household school capital. However, the lack of a scale of educational levels common to the three countries proved to be a serious shortcoming. Indeed, while in Argentina and Chile the specific item responses were based on a scale that distinguished university degrees from

other "tertiary" education (known in Argentina as "non-university higher education" and as "technical" education in Chile), the questionnaire applied in Brazil only made available the "*Ensino superior* [Higher education]" category, encompassing all forms of postsecondary education. But given the characteristics of the population under scrutiny, consisting of university-level researchers, giving up the distinction between university and other types of tertiary education seemed a ruinous alternative[16] because it would largely eliminate the differences in the cultural capital of origin and its possible effects on modes of acquisition and degrees of use of English and other foreign languages. Note that previous research has shown that, as far as the skills in foreign languages are concerned, this distinction within the population with higher education is quite relevant: for instance, the Argentinian ENES-PISAC survey[17] established that 80% of university students knew at least one foreign language in 2015, compared to about 55% of those studying in other types of tertiary education. Moreover, the proportion of students highly qualified in foreign languages, and even bilingual, was also much higher among the former stratum.

The solution adopted in this research consisted of developing an ordinal scale ranging from zero to seven educational levels[18] and of retrieving at least basic information about Brazilian researchers' parental occupation to correct the score for those who worked in university professions (physicians, lawyers, engineers, university professors, etc.), assigning them the corresponding educational score for the university level.

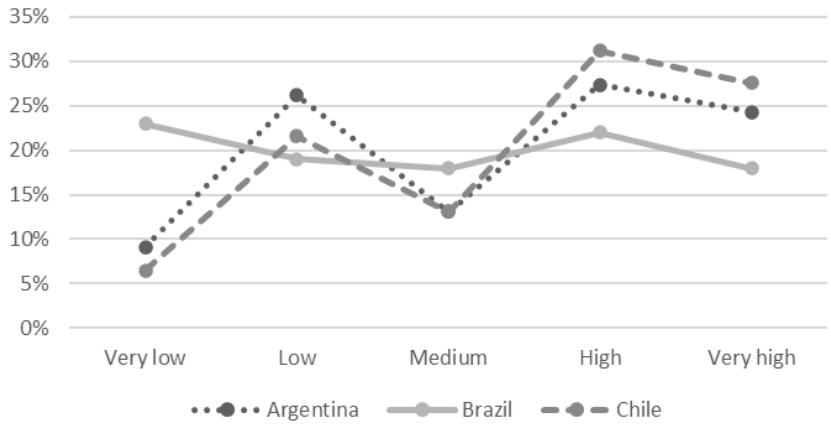
However, the comparison of researchers' social origins (taking parents' educational attainment as a proxy) and the school capital of the parental household (based on the both parents' educational attainment) must not be mechanically based on the face value of indicators due to difficulties in interpreting them. Even if we assume that the operational definitions of these indicators functioned efficiently (i.e., all three samples uniformly understood the offered questions and response categories met the requirements of completeness and mutual exclusivity), the face value of those indicators, if taken uncritically, can result in significant biases. When comparing them, one has to consider both the cross-national (comparing the distributions among countries) and intra-national levels (comparing the distribution of the sample in relation to the general population of each country).

Chilean researchers have (or had) parents with higher educational levels since not only is the proportion of mothers and fathers with university studies higher, but it is also lower that of those who only completed, at most, primary education. In contrast, Brazilian researchers are those who generally have (or had) parents with lower formal educational attainment with significantly higher percentages of mothers and fathers who only completed primary education at most and with lower proportions of

fathers (and especially mothers), who completed university studies. However, it must be considered that the Brazilian sample is the oldest and, unsurprisingly, their parents were on average born before those of their Chilean and Argentinian counterparts. This is significant because the educational attainment of populations has increased from generation to generation. This could also explain the largest gap in the Brazilian sample of the percentage of mothers and fathers with university studies since females have caught up with men's educational attainment in recent decades due to the framework of a more complex process of societal changes that evolved over time. Chilean data must also be read in the light of these historical processes: therefore, the high educational attainment of their parents, especially in the case of older researchers (compared to Argentinian ones), strongly indicates social selectivity in access to the scientific system. In the case of Argentina, parents' educational attainment lie above its national average for the same generations, but the fact that those levels are generally high is also due, at least in part, by the greater weight of young researchers in its sample.

Analyzing not only mothers and fathers' educational attainment separately but also the school capital of households shows similar patterns for Argentinian and Chilean researchers, as in Fig. 1. The percentage of researchers who come from households with very low school capital is limited, especially in the Chilean case, whereas the percentages of those with high or very high school capital are significant, and even more so among Chilean researchers. In the Brazilian sample, paternal household school capitals oscillated at all levels between 17.5% and 23.5%, the latter corresponding to researchers raised in households with very low school capital.

Figure 1 – Researchers' parental household school capital by country



Source: Project NEIES Mercosur # 3/2015/ECAPIN Survey.

When comparing these data on parental household's school capital as well as on mothers and fathers' educational attainment, the characteristics of the social structure and sociohistorical context of each country should also be regarded. In the case of Argentina, it is important to consider, on the one hand, the relative youth of the sample — as already noted — and, on the other, the fact that, by Latin American standards, the educational levels of the population have been traditionally high, and comparatively even more so in the past. Therefore, although this sample over-represents households with high school capital, in relation to their weight in the general population, the selectivity of the scientific system seems to be more attenuated than in the Chilean case, in which the over-representation of parental households with high school capital is very extreme, especially given the older age of these researchers in comparison with Argentinian ones. In the Brazilian case, on the other hand, it is also important to consider researchers' age since many of them were born at a time in which education was less widespread. This explains, at least in part, the largest proportions of researchers in this country with parents with lower educational attainment. However, data also provide evidence of a dynamic society, with a strong upward social mobility, especially in the context of the processes of industrialization and economic expansion that, even with abrupt ups and downs, characterized this country from the second half of the 20th century onward.

In short, if Argentina is compared to Brazil, although the cultural capitals of origin are clearly higher in the former, this cannot be taken as an uncontested indicator that they belong to higher social classes. In other words, for a meaningful comparison, indicators require an interpretation in the light of the characteristics of each sample — for instance, age distribution — and the social structure of each country.

Finally, and regarding the subject addressed above, we find another issue of great importance: what do indicators mean, or rather, what do they indicate? In the methodological literature on cross-national surveys, much of the discussion about the equivalence of indicators has revolved around aspects arguably more linked to reliability: how to produce operational definitions — considering language specificities — that can be understood uniformly in different contexts of comparison and how to generate equivalent response categories across contexts. The equivalence of meaning of complex theoretical concepts — such as democracy, freedom, authoritarianism, etc. — has also been discussed and empirically investigated.

Less attention has been given to the problem of validity since all too often researchers tend to assume the "universal" validity of indicators. However, it should be remembered, as Marradi (2018) points out, that the shared semantic content of indicators with the concepts they indicate is always partial[19]. In this sense, every indicator has an *indicating component* and an *"alien" component* (the latter eventually indicating other conceptual

properties). To illustrate this, we can turn to two examples drawn from the research at stake.

The scientific productivity of Chile — according to the standard indicator used internationally — is much higher than that of Argentina and Brazil. The well-established indicator on which this statement is based is valid but it is also arguable that it indicates phenomena (the *alien component*) other than productivity. For instance, Chile also has a higher level of internationalization, with the lowest proportion of researchers with local PhDs, and the highest (by far) of researchers with PhDs obtained in English-speaking countries: more than 31%, compared to just over 5% for Argentinians and 17.4% for Brazilians. Thus, the productivity indicator, to the extent that it is based on mainstream publications, also indicates a greater adherence to international standards of academic production and publication — with a predominance of articles in English published in mainstream journals (Beigel, 2014) —, as well as the impact of a scientific policy that encourages this type of publication. Moreover, the greater proportion of publications of this type also relates to the larger number of local journals indexed in the mainstream circuit and the higher percentage of doctors trained in Anglo-Saxon universities. Note that previous research has shown that, in the case of peripheral and semi-peripheral scientific contexts — as those under study —, researchers with PhDs from American, British, and other English-speaking countries universities are more prone to publish in English in mainstream journals (Calvo et al., 2019; Gantman, 2011). In the Argentinian case, on the contrary, its much lower scientific productivity, as measured by the traditional indicator, also accounts for the more endogenous and self-referential nature of its scientific system. In fact, it is the one with the highest proportion of researchers born in the country (97%) — a figure that drops to 89% for Brazilians and 81% for Chileans — and the highest percentage of researchers with local PhDs (84%). On the other hand, Argentina has had significant movements meant to defend the relevance of publishing in Spanish as well as the importance of books, which, in turn, have reinforced "heterodox" production and publication patterns in this academic community, particularly in the case of the social sciences and the humanities (Piovani, 2018; 2019). Therefore, if we considered *all* scientific production — which could be quite difficult given that it should also include books and articles published in non-indexed journals —, the difference in productivity between countries would probably be much smaller. In short, the standard productivity indicator undoubtedly shows something relevant about scientific output, but it also accounts for certain characteristics of local scientific systems, researchers' educational trajectories, their degree of adherence to international standards of production and publication, and local scientific policies and academic production patterns. If these aspects are neglected, comparisons across countries, based solely on the

face value of the productivity indicator, can bias or distort representations of the subject in question.

The second example deals with a paradox: although the Argentinian population at large has the highest proficiency in English of the three countries, according to the English Proficiency Index (EPI)[20], its scientists are those who least use English for academic purposes. And while its entire population's skills in the English language have increased from generation to generation, among the younger scientists of this country, these skills seem to be lower than among the older ones. What, though, does this indicator, that seems to make little sense in the case of Argentinian scientists, tell us? In general, it could be argued that it hints to the so-called process of de-elitization of higher education (Benza & Kessler, 2020), about which much has been written in recent years in Latin America. And since the gap in linguistic skills — in the English language — amidst Argentinian scientists and the general national population is less significant than in Brazil and Chile, this indicator may also suggest lower social selectivity in the access to the scientific system in Argentina since as Kaplan and Piovani (2018) have shown, these skills are strongly associated with social class and parents' educational attainment.

CONCLUDING REMARKS

In this study, we aimed, rather than describing the results of an investigation on linguistic skills of South Americans scientists in their own right — which would have required far more elaborate and comprehensive analyses of its data, to take them as examples of some key problems of comparability and equivalence in cross-national research.

On the one hand, we dealt with the limitations that arise when trying to define exactly equivalent populations of scientists given different local scientific traditions and institutions. Moreover, and even if we pragmatically achieved a reasonable equivalence of target populations, we showed how the nature and characteristics of samples can undermine comparability. In the case of the research referenced in this study, the self-selected samples of scientists for each country, which by no means can be considered random, present serious limitations and biases for establishing comparisons and producing sound statistical inferences.

On the other hand, we have analyzed the problem of equivalence of empirical indicators. In this sense, we have argued that this issue, which Van Deth (2009: 95) rightly considers as "one of the nastiest problems facing comparativists," goes well beyond the usually acknowledged matters of shared meaning and adequate translation of operational definitions. This implies that it is not just a matter of measurement and reliability.

To ensure equivalence when comparing populations from different social and cultural settings, indicators (and the empirical data linked to them)

must also be considered in the light of a complex set of aspects that demand a thorough familiarity with each particular context (historical, sociological, political, linguistic, etc.), as well as the obvious expertise in the concrete field of research.

Achieving comparability and equivalence is not a technical issue that can be solely — and uncritically — solved on the basis of impersonal — objectified — statistical knowledge. Instead, and acknowledging the enormous value of statistical techniques, it demands artisan-like work specific to any given piece of research, likely to rely — at least in part — on personal and tacit knowledge (Polanyi, 1958; 1966).

Finally, it seems appropriate to point out that, as with any methodological subject, the questions of comparability and equivalence can be understood within a complex space of critical analysis, which Gallino (1978) and Marradi (1996) have characterized as a *continuum* defined by epistemological and technical poles. Hence, solving questions of comparability and equivalence (from a methodological point of view) requires acknowledging, as Bruschi (1991) points out, that the disregard of the technical pole carries the risk of reducing research practices to mere abstract speculation incapable of guiding concrete investigative paths, whereas the annulment of the epistemological pole tends to give way to a ritualistic and uncritical use of the techniques, disconnected from the ontological and epistemological foundations of scientific knowledge.

Received on 27-mar-2022 | Revised on 23-feb-2023 | Approved on 17-mar-2023

Juan Ignacio Piovani is PhD in Social Science Methodology (Sapienza University of Rome, Italy). Full Professor at the Faculty of Humanities of the National University of La Plata (Argentina) and principal researcher of CONICET and MECILA. His research has addressed issues related to the history of sociological research methods and techniques, to the field of social sciences in Argentina, and to the labor market and social inequalities. He is the author of several books and numerous articles published in European and Latin American journals.

Denis Baranger is PhD in Social Sciences (University of Buenos Aires, Argentina). Professor emeritus at the National University of Misiones (Argentina), with teaching and research functions in the Graduate Program in Social Anthropology. Specialized in methodology and epistemology, he works on the reception of Bourdieu's work in Argentina and Latin America and on the Argentine field of social sciences. He is the author of *Epistemología y metodología en la obra de Bourdieu* (Buenos Aires: Prometeo, 2004) and other books, as well as numerous articles and chapters published in Argentina and abroad.

NOTES

1 Project NEIES Mercosur # 3/2015: Academic internationalization in the Southern Cone. Comparative study of language skills of academics from selected universities in Chile, Brazil and Argentina.

2 ECAPIN: Encuesta de Capacidades Lingüísticas e Internacionalización [Internationalization & Language Abilities Survey].

3 Note that, although widespread in the language of the social sciences, the terms "measure" and "measurement" entails serious difficulties and limitations. For an appraisal of the debate concerning measurement in the social sciences, see Marradi (1981).

4 Cited by Ragin (2014).

5 Cited by Fideli (1998).

6 The expression "trans-contextual" also applies to the comparison of the same object at two distant points in time since it is assumed that the passage of time may result in a significant change of context. Therefore, trans-contextual investigations can be, in turn, cross--sectional or cross-temporal.

7 This gives rise to the differentiation between synchronic and diachronic comparison.

8 Cited by Bynner y Chisholm (1998)

9 Bourdieu was aware of the problems in cross-national studies since his research in the mid-60s on European art museums was conducted in France, Greece, Netherland, Poland, and Spain. At that time, he stated that "It was impossible to ignore the fact that to attempt a formal homogeneity of the codes was to run the risk, inherent in every comparison of abstract and falsely interchangeable evidence, of comparing facts which are formally comparable but actually incomparable and, conversely, of omitting to compare formally incomparable facts which are really comparable." (Bourdieu & Darbel, 1991: 12)

10 In the following paragraphs, we borrow from Baranger (2005: 108-112).

11 A few years before, Bourdieu and Boltanski questioned themselves regarding social mobility surveys on the problem of "the permanence of the relationship between words and things, between qualifications and positions,

between the nominal and the real: what is the sense of identifying the 1880 elementary school teacher with the 1930 teacher and the 1974 teacher? Is the son of an elementary school teacher really the son of a teacher in the sense that he is a teacher himself?" (1975: 95).

12 Abbott, examining standard positivist articles concluded that, in these, "variables do the acting," adding that "the realistic metaphysics implicit in treating variables (universals) as agents was last taken seriously in the age of Aquinas", even though sociological positivism usually describes itself as nominalist (Abbott, 1992: 54-58).

13 This explains Bourdieu's choice of Multiple Correspondence Analysis (MCA), in which each indicator (each modality from each variable) will be located in the space of properties in a position of differential distance-proximity from the others.

14 Since January 1, 2020, CONICYT has been relabeled as National Research and Development Agency (ANID).

15 This is well-known in the methodological literature: it was pointed out long ago by Blumer (1956) and Cicourel (1964).

16 In Argentina, for example, the training of teachers for primary education (and often that of teachers for secondary education) is still carried out in tertiary establishments. Therefore, to analyze the role played by cultural capital, it was crucial to distinguish between tertiary and university education: usually, a primary teacher is not equivalent to a university professor, and a surgical assistant is not equivalent to a surgeon.

17 https://www.argentina.gob.ar/ciencia/pisac/bases-de-datos.

18 0: Without formal education; 0.5: Incomplete primary education; 1: Complete primary education; 2: Incomplete secondary education; 3: Complete secondary education; 4: Incomplete tertiary (Chile: technical) education; 5: Complete tertiary (Ch: technical) education; 5: Incomplete university education; 6: Complete university education (undergraduate); 7: Postgraduate education.

19 This idea was originally formulated by Lazarsfeld (1958: 104): "by their nature the indicators are many and their relation to outside variables are usually both weaker and more unstable that the underlying characteristic which we would like to measure. To put it in more formal

language, each individual indicator has only a probability relation to what we want to know."

20 The world's... (2022).

REFERENCES

Abbott, Andrew. (1992). What do cases do? Some notes on Activity in Sociological Analysis. In: Ragin, Charles & Becker, Howard (orgs.). *What Is a Case? Exploring the Foundations of Social Inquiry*. Cambridge: Cambridge University Press, p. 53-82

Almeida, Ana Maria et al. (2022) Social origin, language skills and scientifc publication patterns in Argentina, Brazil and Chile. *Tempo Social,* 34/3, p. 75-99.

Almond, Gabriel. (1966). Political Theory and Political Science. *American Political Science Review*, 60, p. 877-78.

Apter, David E. (1971). Comparative studies: A Review with some Projections. In: Vallier, Ivan (org.). *Comparative Methods in Sociology. Essays on Trends and Applications*. Berkeley: University of California Press, p. 3-15.

Archenti, Nélida & Piovani, Juan Ignacio. (2018). La Comparación. In: Marradi, Alberto et al. (orgs.). *Manual de Metodología de las Ciencias Sociales*. Buenos Aires: Siglo Veintiuno, p. 305-324.

Ariely, Gal & Davidov, Eldad. (2012). Assessment of Measurement Equivalence with Cross-National and Longitudinal Surveys in Political Science. *European Political Science*, 11, p. 363-377.

Baranger, Denis. (2005). *Epistemología y metodología en la obra de Pierre Bourdieu*. Buenos Aires: Prometeo.

Beigel, Fernanda. (2014). Introduction: Current tensions and trends in the World Scientific System. *Current Sociology*, 62/5, p. 617-625.

Benza, Gabriela & Kessler, Gabriel. (2020). *Uneven Trajectories. Latin American Societies in the Twenty-First Century*. Cambridge: Cambridge University Press.

Blasius, Jörg & Thiessen, Victor. (2006). Assessing Data Quality and Construct Comparability in Cross-National Surveys. *European Sociological Review*, 22/3, p. 229-242.

Blumer, Herbert. (1956). Sociological Analysis and the 'Variable.' *American Sociological Review*, 2/6, p. 683-90.

Bourdieu Pierre & Boltanski, Luc. (1975). Le titre et le poste: rapports entre le système de production et le système de reproduction. *Actes de la recherche en sciences sociales*, 1-2, p. 95-107.

Bourdieu, Pierre. (1984). *Distinction. A Social Critic of the Judgement of Taste*. Cambridge: Harvard University Press.

Bourdieu, Pierre & Darbel, Alain. (1991). *The Love of Art, European Art Museums and their Public*. Cambridge: Polity Press.

Brady, Henry E. (1985). The Perils of Survey Research: Inter-Personally Incomparable Responses. *Political Methodology*, 11/3-4, p. 269-290.

Brick, J. Michael. (2011). The future of survey sampling. *Public Opinion Quarterly*, 75/5, p. 872-888.

Bruschi, Alessandro. (1991). Logica e metodología. *Sociologia e Ricerca Sociale*, 35, p. 30-55.

Bynner, John & Chisholm, Lynne. (1998). Comparative Youth Transition Research: methods, meanings and research relations. *European Sociological Review*, 14/2, p. 131-150.

Calvo, Ernesto et al. (2019). Investigando las influencias internacionales en las ciencias sociales argentinas. *Revista Latinoamericana de Metodología de las Ciencias Sociales*, 9/2.

Cicourel, Aaron V. (1964). *Method and Measurement in Sociology*. New York/ London: The Free Press/Collier-Macmillan.

Converse, Jean & Presser, Santley. (1986). *Survey questions: Handcrafting the standardized questionnaire*. Thousand Oaks: Sage.

Davidov, Eldad & De Beuckelaer, Alain. (2010). Testing the equivalence of an instrument to assess Schwartz' human values: How harmful are translations? *International Journal of Public Opinion Research*, 22/4, p. 485-510.

Davidov, Eldad et al. (2014). Measurement equivalence in cross-national research. *Annual Review of Sociology*, 40, p. 55-75.

Deutscher, Irwin. (1968). Asking questions cross-culturally: Some issues of linguistic comparability. In: Becker, Howard et al. (orgs.). *Institutions and the person*. Chicago: Aldine, p. 318-341.

Eisenstadt, Samuel N. (1968). Social Institutions: Comparative Study. In: Sills, David L. (org.). *International Encyclopedia of the Social Sciences*. New York: Macmillan and Free Press, p. 409-429.

Fideli, Roberto. (1998). *La Comparazione*. Milan: Franco Angeli.

Gallino, Luciano. (1978). *Dizionario di Sociologia*. Turin: UTET.

Gantman, Erneto R. (2011). La productividad científica argentina en ciencias sociales: economía, psicología, sociología y ciencia política en el CONICET (2004-2008). *Revista Española de Documentación Científica*, 34/3, p. 408-425.

George, Janet E. G. (1986). Comparative Social Research: Issues in Comparability. *Acta Sociologica*, 29/2, p. 167-170.

Hantrais, Linda. (2009). *International Comparative Research. Theory, Methods and Practice*. London: Palgrave Macmillan.

Harkness, Janet A. et al. (2010). *Survey Methods in Multicultural, Multinational, and Multiregional Contexts*. New York: Wiley.

Heen, Miliaikeala et al. (2011). A Comparison of Different Online Sampling Approaches for Generating National Samples. Research in Brief. Las Vegas: Center for Crime and Justice Policy, University of Nevada. Available from: https://www.unlv.edu/sites/default/files/page_files/27/ComparisonDifferentOnlineSampling.pdf. Accessed on: 29 June 2020.

Jowell, Roger. (1998). How Comparative Is Comparative Research? *American Behavioral Scientist*, 42/2, p. 168-177.

Kaplan, Carina & Piovani, Juan Ignacio. (2018). Trayectorias y capitales socioeducativos. In: Piovani, Juan Ignacio & Salvia, Agustín (orgs.). *La Argentina en el siglo XXI: cómo somos, vivimos y convivimos en una sociedad desigual*. Buenos Aires: Siglo Veintiuno, p. 221-263.

King, Gary et al. (2003). Enhancing the Validity and Cross--Cultural Comparability of Survey Research. *American Political Science Review*, 97/4, p. 567-583.

Lasswell, Harold. (1968). The Future of the Comparative Method. *Comparative Politics*, 1, p. 3-18.

Lazarsfeld, Paul F. (1958). Evidence and Inference in Social Research. *Daedalus*, 87/4, p. 99-130.

Lee, Sunghee & Valliant, Richard. (2009). Estimation for Volunteer Panel Web Surveys Using Propensity Score Adjustment and Calibration Adjustment. *Sociological Methods and Research*, 37, p. 319-43.

Lijphart, A. (1971). Comparative Politics and the Comparative Method. *American Political Science Review*, 65/3, p. 682-693.

Lijphart, Arend. (1975). The Comparable-Cases Strategy in Comparative Research. *Comparative Political Studies*, 8/2, p. 132-157.

Loosveldt, Geert & Sonck, Nathalie. (2008). An evaluation of the weighting procedures for an online access panel survey. *Survey Research Methods*, 2, p. 93-105.

Marradi, Alberto. (1981). Misurazione e scale Qualche riflessione e una proposta. *Quaderni di Sociologia*, 21, p. 117-132.

Marradi, Alberto. (1991). Comparación. In: Reyes, Ramón (org.). *Terminología Científico-Social*. Barcelona: Antropos, p. 65-84.

Marradi, Alberto. (1996). Metodo come arte. *Quaderni di Sociologia*, 40, p. 71-92.

Marradi, Alberto. (2018). Indicadores, validez, construcción de índices. In Marradi, Alberto et al. (orgs.). *Manual de Metodología de las Ciencias Sociales*. Buenos Aires. Siglo Veintiuno, p. 195-231.

Miller, Joanne et al. (1981). Assessing Comparability of Measurement in Cross-National Research: Authoritarian--Conservatism in Different Sociocultural Settings. *Social Psychology Quarterly*, 44/3, p. 178-191.

Mullen, Michael R. (1995). Diagnosing measurement equivalence in cross-national research. *Journal of International Business Studies*, 26, p. 573-596.

Pinto, Louis. (1999). *Pierre Bourdieu et la théorie du monde social*. Paris: Albin Michel.

Piovani, Juan Ignacio. (2001). Los estudios comparativos: estrategias metodológicas para la investigación empírica en relaciones internacionales. *Relaciones Internacionales*, X/20, p. 97-108.

Piovani, Juan Ignacio. (2006). *Alle origini della statistica moderna*. Milan: Franco Angeli.

Piovani, Juan Ignacio. (2018). Estilos de producción en el campo de las ciencias sociales en Argentina. *Cuadernos de Información y Comunicación,* 23, p. 125-141.

Piovani, Juan Ignacio (2019). Styles of academic production in the Argentine social sciences: Heterogeneity and heterodoxy. *Serendipities. Journal for the Sociology and History of the Social Sciences*, 4/1-2, p. 27-48.

Polanyi, Michael. (1958). *Personal Knowledge: Towards a post--critical philosophy*. London: Routledge and Kegan Paul.

Polanyi, Michael. (1966). *The tacit dimension*. New York: Doubleday.

Ragin, Charles. (2014). *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley: University of California Press.

Schonlau, Mathias et al. (2007). Are 'Webographic' or Attitudinal Questions Useful for Adjusting Estimates from Web Surveys Using Propensity Scoring? *Rand Working Paper Series* WR, 506. Available from: https://www.rand.org/content/dam/rand/pubs/working_papers/2007/RAND_WR506.pdf. Accessed on: 29 June 2020.

van Deth, Jan W. (1998). *Comparative Politics. The Problem of Equivalence*. London: Routledge.

van Deth, Jan W. (2009). Establishing Equivalence. In: Landman, Todd & Robinson, Neil (orgs.). *The SAGE Handbook of Comparative Politics*. London: Sage, p. 84-100.

Winch, Peter. (1964). Understanding a primitive society. *American Philosophical Quarterly*, 1/4, p. 307-324.

The world's largest ranking of countries and regions by English skills. *EF*. Available on: https://www.ef.com/wwen/epi/. Accessed on: 31 mar 2023.

COMPARABILITY AND EQUIVALENCE IN SOCIAL RESEARCH

Keywords
Comparative research;
Comparability;
Equivalence;
Cross-national studies.

Abstract

This study discusses some problematic issues around comparability and equivalence in empirical research, particularly focusing on the case of indicators and taking a cross-national study on linguistic skills and use of foreign languages in the scientific production of researchers from Argentina, Brazil, and Chile as reference. The discussion is framed within broader epistemological and methodological debates concerning the status of comparisons in the social sciences and, mainly, the case of cross-national comparative studies. Within these debates, and resorting to concrete examples arising from research into scientists' linguistics skills, we address the comparability issue of target populations and samples, as well as that of indicator equivalence, with emphasis in the problems of questionnaire design, reliability, and validity.

COMPARABILIDADE E EQUIVALÊNCIA NA PESQUISA SOCIAL

Palavras-chave
Pesquisa comparativa;
Comparabilidade;
Equivalência;
Estudos transnacionais.

Resumo

Este artigo discute algumas questões problemáticas em torno da comparabilidade e equivalência em pesquisas empíricas, focando particularmente no caso dos indicadores, e tomando como referência um estudo transnacional sobre habilidades linguísticas e uso de línguas estrangeiras na produção científica por pesquisadores da Argentina, Brasil e Chile. A discussão se enquadra em debates epistemológicos e metodológicos mais amplos sobre o status da comparação nas ciências sociais e, principalmente, no caso dos estudos comparativos transnacionais. No âmbito destes debates, e recorrendo a exemplos concretos decorrentes da investigação das competências linguísticas dos cientistas, é abordada a questão da comparabilidade de populações-alvo e amostras, bem como da equivalência de indicadores, com destaque para os problemas de desenho de questionários, confiabilidade e validade.