

# METHODOLOGICAL FORMS AND CUSTOMS IN SPANISH PSYCHOLOGY: AN ANALYSIS THROUGH THE LIFE OF *PSICOTHEMA* (1990-1999)

Ignacio Montero and Orfelio G. León  
*Universidad Autónoma, Madrid*

*The present paper explores in a descriptive way the kind of research methods most frequently used by Spanish researchers in psychology. The production of *Psicothema* (1990-1999) was analyzed as illustrative of work in this discipline in Spain. A coding system was developed for the categorization of different types of research designs. 'Study' rather than 'paper' was considered as the unit of analysis, given the fact that some papers included several studies. Usage proportions of eight types of studies are illustrated: theoretical, experimental, quasi-experimental, observational descriptions, survey studies, instrumental studies, case studies and one-subject experimental designs. These proportions are presented in relation to year of publication. Finally, we discuss the implications for teaching of issues such as methodological variability and errors in the interpretation of interactions.*

*El presente trabajo explora de forma descriptiva el tipo de metodologías que los investigadores de la Psicología española usan más frecuentemente. Para ello se ha analizado la producción de la revista *Psicothema* (1990-1999) a modo de ejemplo. Se ha elaborado un sistema de categorización de los diferentes tipos de diseños tomando como unidad de análisis cada estudio, independientemente del número de estudios que contuviera cada artículo. Se ilustran las proporciones de uso de ocho tipos de trabajos: teóricos, experimentales, cuasi experimentales, descriptivos mediante observación, descriptivos mediante encuestas, instrumentales, descriptivos de casos y experimentos de caso único. Dichas proporciones se analizan en función de su evolución a lo largo de la década. Se finaliza discutiendo las implicaciones que, para la enseñanza, pudieran tener la variabilidad metodológica y la presencia de errores en la interpretación de los diseños complejos con interacciones significativas.*

A part of the authors' work being to teach the rudiments of methodology to future psychologists – perhaps future researchers, who knows? – it occurred to them to take a systematic and replicable look at the recent production of researchers in psychology in Spain. Given the boom in publication over the last fifteen years in our country, it has been necessary to use some kind of heuristic bias – perhaps that of accessibility – that would permit us to accomplish a task that would otherwise have been impossible with the available resources. This bias was the coincidence with the ten years of life of the journal *Psicothema*, which, in addition to being well regarded among researchers and having a high degree of impact, has the peculiarity of publishing work in all fields of psychology, which is advantageous for the objective of this study.

This objective, then, is to present a description of the

methodologies, the research designs, used in the issues of *Psicothema* over the ten-year period from 1990 to 1999, in the knowledge that, despite being a biased sample, it may be illustrative of the forms and customs referred to in the title of this work (for an analysis dealing more directly with the content of the research published in this journal, the reader should consult Moreno and Sánchez, 1998). Given that our objective is of a descriptive nature, we use the variable time – calendar year – as a basic organizer of this description, which we carry out using the terminology employed in a textbook recently published in our language (León and Montero, 1997; see Kerlinger and Lee, 2000; Shaughnessy, Zechmeister and Zechmeister, 2000, as recent manuals in English).

## METHOD

### *Unit of analysis*

The unit of analysis we decided to use for this documentary research was "study", considering as independent units each one of the studies published within a single article. Review articles were considered as theoretical studies. We did not include publications presented within the section "methods, programmes and instruments."

---

The original Spanish version of this paper has been previously published in *Psicothema*, 2001, Vol. 13. No 4, 671-677

.....  
*Correspondence concerning this article should be addressed to Ignacio Montero García-Celay, Facultad de Psicología, Universidad Autónoma de Madrid. 28049 Madrid. Spain.*

E-mail: nacho.montero@uam.es.

## Materials

In order to carry out the descriptive analysis we designed a system of codification, derived largely – as we pointed out earlier – from the terminology of León and Montero (1997). The system was made up of eight main categories, each of which contained several subcategories. The eight main ones, in addition to that of theoretical studies already mentioned, were descriptive studies by means of observation, descriptive studies by means of surveys, experimental studies, quasi-experimental studies, single-subject experiments, instrumental studies (development of tests and equipment, design and/or adaptation of these) and descriptive case studies. Appendix A includes a definition of each category and its subcategories.

Reliability of the categorization system was studied by calculating the percentage of agreement between coders. An advanced psychology student was trained to use the coding system and as a sample we took, at random, all the studies published in a particular year. With these studies categorized independently by the student and one of the authors, an agreement level of 95% was obtained. In cases of disagreement, the inclusion criteria were clarified until agreement was reached. On classifying the publications from another year, an agreement of 99% was achieved.

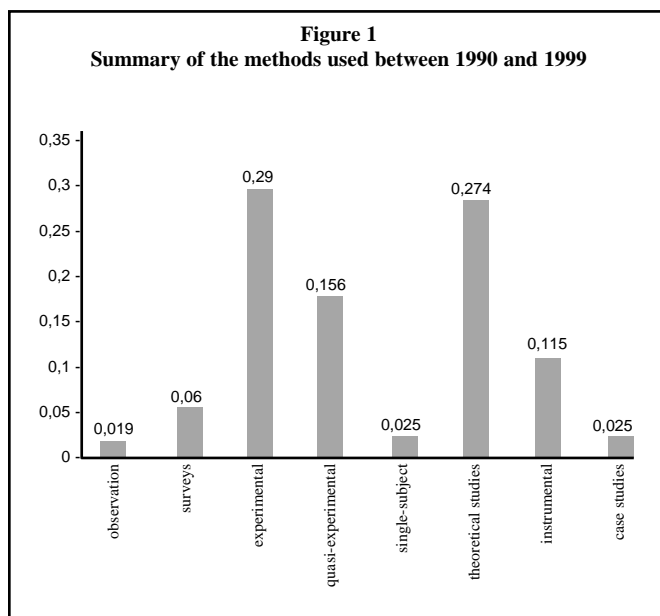
## Design and procedure

This work is a descriptive study by means of the analysis of documents. The description was carried out using a system for classifying the methodology employed in each of the studies analyzed. Once this system had been developed, its reliability was tested using the procedure referred to in the previous section. Furthermore, we used the variable time as the axis of the description, the calendar year of publication of each study constituting the form of categorizing this variable.

## RESULTS AND DISCUSSION

Table 1 shows the distribution of the number of studies considered for our research, grouped according to year of publication. The marked increase from 1993 to 1994 is due to the fact that the journal went from publishing two numbers per year to publishing three. We should like to stress that the studies in the section “methods, programmes and instruments” were not taken into account. Total number of articles published in the decade studied was 367. This indicated an annual mean of almost 37, if we take the decade as a whole. If we separate the publications in two periods according to the number of issues per year, we find that the first period (1990-93) has an annual mean of 23 articles published, 11.5 per number, while the second period (1994-99) shows a mean of almost 45.8 studies each year, a mean of over 15.3 per issue. This implies an increase in the number of studies published that is independent of the increase in number of issues: the mean of 23 articles with two numbers almost doubles to reach 45.8 with three issues per year.

With regard to the type of methodology used in each one of the studies, Figure 1 shows a histogram of the distribution across the decade. The proportion of reviews (theoretical studies) is 0.274, more than a quarter of the total. Almost another third of the publications is accounted for by experimental studies (0.29), though it must be borne in mind that this group includes studies that analyze variables not always manipulated by the researcher. It was sufficient for one variable to be manipulated for the study to be included in this category. The next group of studies as regards frequency is that which includes those labelled quasi-experimental, which account for 0.156 of the total, a little over half the share represented by the previous group. This group is followed by those studies



**Table 1**  
Distribution of units of analysis by year of publication

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	Total
Studies	18	22	30	22	33	41	47	49	54	51	367

we call instrumental, and which involve the development of tests and equipment and their design and/or adaptation. The proportion in this case is 0.115. Descriptive studies by means of surveys account for 0.06 of the total, half as much as the instrumental studies and less than a quarter of the proportion represented by experimental studies. Single-subject experimental studies and case studies each represent proportions of 0.025, while observational descriptions occupy the smallest proportion, at 0.019. We find, therefore, that approximately 42% of the publications with empirical content are experimental in nature, and almost 17% are instrumental studies. Between them, these two types of study account for almost two-thirds of the articles with empirical content.

In the following figure we try to illustrate the evolution of the different types of study. Figure 2a shows the changes in the number of theoretical studies in relation to the experimental articles. The theoretical works initially account for a very high proportion – 0.55 in 1990 – but this proportion falls to between 0.10 and 0.15 in the last two years studied. In contrast, the experimental studies are poorly represented in the first three years, with 0.10 in 1992, but consolidate themselves at a proportion of 0.35 or over from 1996. The evolutions of these two types of study cross around 1995, at which point they each account for a proportion of around 0.25.

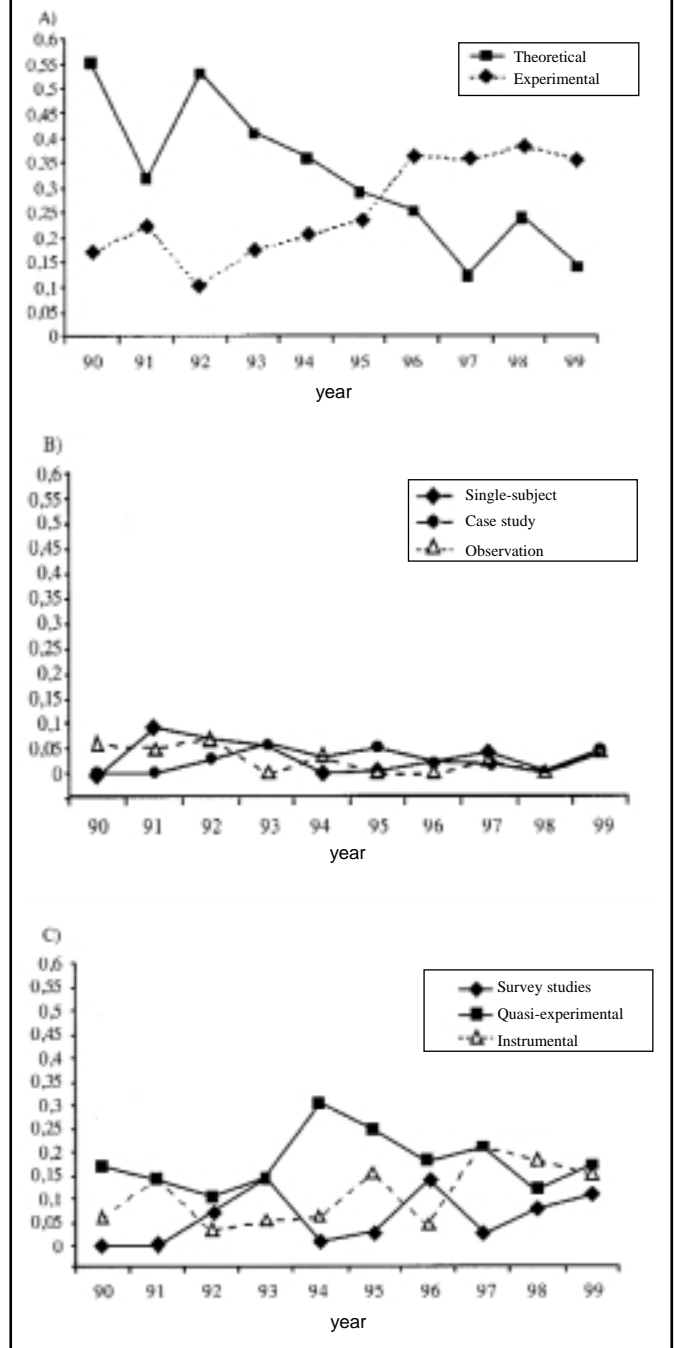
Figure 2b shows the evolution of the observational descriptions, the descriptive case studies and the single-subject experiments. These three types of study have in common that they account for low proportions – always less than 0.1 – and show a certain stability throughout the decade.

The quasi-experimental studies, the survey studies and the instrumental studies (see Figure 2c), on the other hand, present a more oscillating profile over the decade. The quasi-experimental works move within a range from 0.1 to 0.3, the survey studies between zero and a proportion close to 0.15 (though their annual mean is 0.06), and the instrumental studies between a minimum proportion of 0.03 and a maximum of 0.20, with a clear difference in favour of the second part of the decade.

Having described the main categories of the coding system employed, we shall now present some data showing how these broad categories are broken down into more detailed ones, bearing in mind that this breakdown can help to better illustrate the comments made up to now. This breakdown affects the experimental and quasi-experimental studies. As far as the first group is concerned, Figure 3 shows the separate evolution of studies with simple experimental designs – of a single independent variable – and factorial designs – with more than one independent variable, at least one being mani-

pulated. As a general comment, we should point out that the crossing of the tendencies in the final year of the analysis may be insignificant, though to be sure of this we would need to continue the data series further. This finding aside, it would appear that, as is to be expected,

**Figure 2 a:**  
Comparative evolution of the frequency of theoretical and empirical studies; **b:** Comparative evolution of the frequency of single-subject experiments, case studies and observation studies; **c:** Comparative evolution of the frequency of survey, quasi-experimental and instrumental studies



there are more studies with more than one independent variable. The relationship between one type of study and the other – ignoring the years in which no simple experiments were published –, goes from parity in 1993 to a ratio of four to one in 1994 and three to one in 1998. We should add that the mean number of independent variables included in the factorial studies (the experimental ones plus the complex prospective “ex post facto” ones) is 2.56, the vast majority being of two (45 studies) or three (26 studies). Eight studies with four independent variables were published, and one with five.

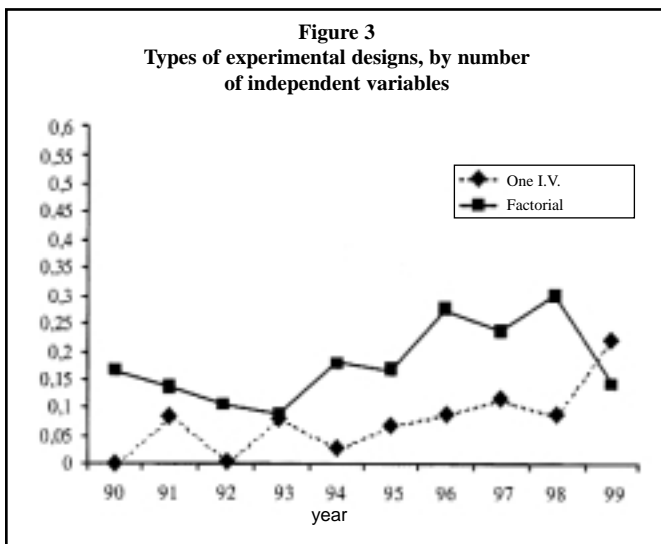
A separate mention should be made of another fact that is not reflected in the figure, but is nevertheless relevant. We are referring to the study of interactions, and the way in which this affects the interpretation of the lower-order effects, be they simple interactions – when there are at least three independent variables – or main effects. Of

the 80 studies analyzed, 46 report the presence of some significant interaction. In 27 of them (58.7%), the interpretability of the lower-order effects may be affected. Lack of sufficient information means that we cannot say for certain that it is affected in all cases, since a significant interaction does not always lead to error in the interpretation of lower-order effects. However, in 15 of these studies we can affirm that the interpretations made of the lower-order effects are incorrect. The errors are of two types. The first type consists in interpreting simple interactions directly once the presence of a higher-order interaction has been detected. To do so is erroneous, because what this interaction means is that at least one of the possible simple interactions does not occur in the same way for all the levels of the variable that is withdrawn from the analysis on passing from one order of interaction to the one immediately below it. The other type of error is simpler to explain. Depending on the type of double interaction produced, the main effects may constitute an inadequate summary of the simple effects. Clearing this up involves studying the simple effects. The cases we mention are those where the effect of an independent variable is reported through the interpretation of its main effect when this is not an adequate summary. This occurs when the influence of the independent variable on the dependent one differs as a function of the levels of the second variable involved in the significant interaction, that is, when qualitatively different simple effects occur (see León and Montero, 2001).

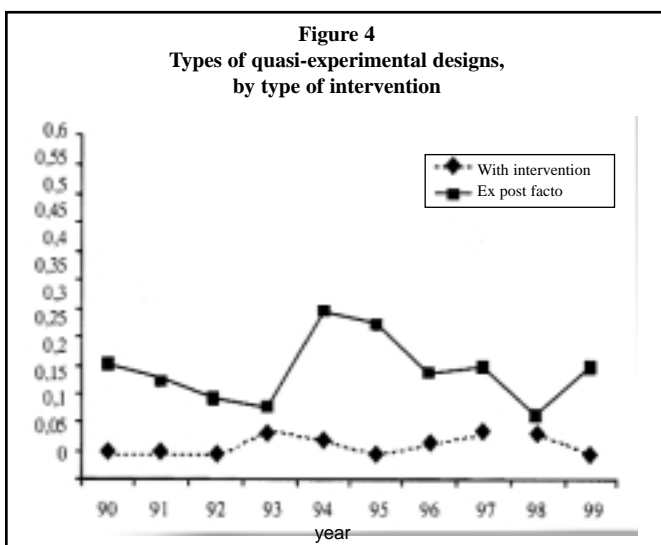
As regards the articles that include studies of a quasi-experimental nature, Figure 4 breaks them down into two main types: designs with intervention and “ex post facto” designs. With a proportion that hovers around 0.15, “ex post facto” designs are always more common than designs with intervention, which remain around a proportion figure of 0.05.

Figure 5 shows that, within the “ex post facto” group, prospective designs (0.72) are more common than retrospective ones (0.25). Of the prospective studies, the most frequently used type is simple (0.33). León and Montero (1997) define a simple prospective design as that which studies a single independent variable that cannot be manipulated by the researcher within a research context in which first of all groups are selected according to their value in the variable under study and subsequently their differences in the dependent variable are studied. In the case of the retrospective studies, the “single group” type is the most numerous (0.14). The manual we have just cited defines a single-subject retrospective design as that carried out in a context in which, within a sample as large and representative as possible, researchers measure simultaneously the dependent variable and all those variables that are candidates

**Figure 3**  
Types of experimental designs, by number of independent variables



**Figure 4**  
Types of quasi-experimental designs, by type of intervention



for explaining, with a minimum of significance, its variability. In this last figure the proportions relate to the total of “ex post facto” type studies.

Another relevant point that arose as a result of our analysis concerns the fact that we found excessive variation in the forms of presentation of the studies, which is surprising, since the journal requires authors to follow the APA norms. As this issue is closely related to the task of reviewers, a series of recommendations for them is presented in Appendix B.

## CONCLUSIONS

Although the nature of this work is purely descriptive and the analysis of a single journal precludes any temptation to generalize, we feel it is possible to make some remarks and raise a few questions by way of conclusion.

First of all, we are concerned about the degree of methodological variability found. We ask ourselves whether it is not somewhat scarce, given the quantity of research methods that are considered as fundamental subjects in psychology teaching curricula. Let us recall that two-thirds of the studies with empirical content were classified within the categories of experimental and instrumental studies. And while we are alarmed about this bias as reflected in the journal, we are even more concerned about its practical implications for our teaching. We feel that the justification of the importance of our disciplines in the training of future psychologists (given that they will become, if not researchers, at least consumers of quality research) is somewhat diminished, if we are to judge by the content of the journal in question. And with regard to the training of future researchers, we are tempted to make the following observation: in optional subjects and doctoral programmes it may not be necessary to explain important methodological developments, in view of the use made of them by those who actually conduct psychological research in our country. Clearly, it is positive to always widen and sharpen students’ knowledge. Frankly, however, the apparent distance between what we teach and what is used amply covers this anticipatory function, at least, we insist, in the light of what we have found in this decade of *Psicothema*.

In addition to the above, another observation we consider to be important concerns the interpretation of lower-order effects when significant interaction effects are found. Although there is a long tradition in the study of errors in data analysis (see Judd, McClelland and Culhane, 1995; Keselman et al., 1998), what we wish to focus on here is related not to the way we *decide* about hypotheses, but to the way we *interpret* the results. At present there is still no clear explanation of the origin of these potential errors of interpretation, which, moreover, can be observed in any Western psychological research

journal (see León and Montero, 2001). What is a fact is that in the analysis presented here, the percentage of cases in which they may appear is high (around 60%). The fact that in a third of all the studies we have found not only the possibility of error occurring but its actual occurrence reflects the relevance of the problem. Although awareness of the problem may be a starting point, an “insight” is not necessarily intrinsically therapeutic.

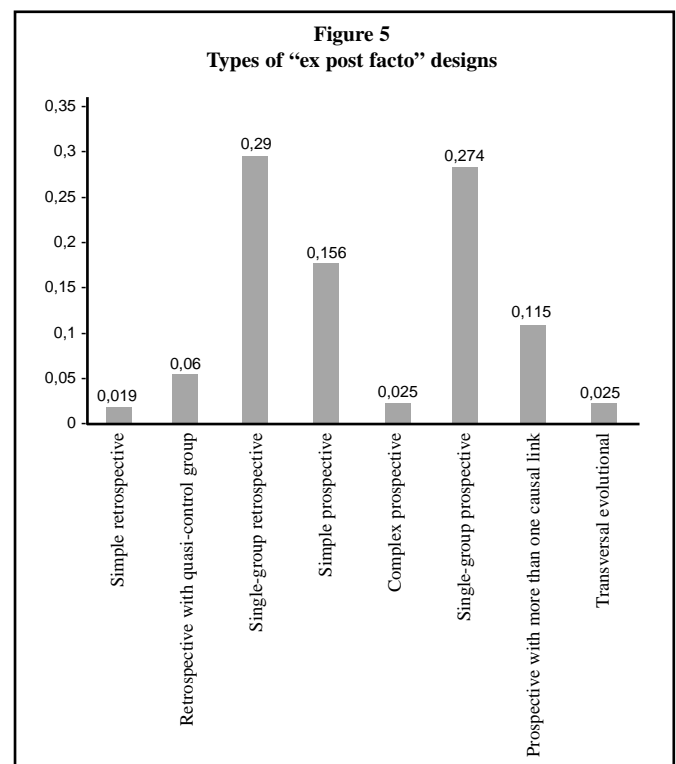
And it is in this way that the two aspects discussed in these conclusions are related. It would seem necessary to call for reflection on the scope of what we teach in our methodological disciplines since, in the light of the present analysis, few different research approaches are used and, on some occasions, their results are interpreted quite incorrectly. One proposal might be the following: “Let us teach less content, but let us teach it better.”

## ACKNOWLEDGEMENTS

The authors of this work would like to thank Beatriz Plans for her excellent work in the classification of the studies analyzed.

## REFERENCES

- A.P.A. (1994). *Publication manual of the American Psychological Association* (4th ed.) Washington, DC: APA.
- Judd, C.M., McClelland, G.H. & Culhane, S.E. (1995). Data analysis: Continuing issues in the everyday



analysis of psychological data. *Annual Review of Psychology*, 46, 433-465.

Keselman, H.J., Huberty, C.J., Lix, L.M., Olejnik, S., Cribbie, R.A., Donohue, B., Kowalchuk, R.K., Lowman, L.L., Petosky, M.D., Keselman, J.C. & Levin, J.R. (1998). Statistical practices of educational researchers: An analysis of their ANOVA, MANOVA and ANCOVA analyses. *Review of Educational Research*, 68, 350-386.

Kerlinger, F.N. & Lee, H.B. (2000). *Foundations of behavioral research* (4th ed.). Fort Worth, TX: Harcourt College Publishers.

León, O.G. & Montero, I. (1997). *Diseño de investiga-*

*ciones. Introducción a la lógica de la investigación en Psicología y Educación* (2ª ed.). Madrid: McGraw-Hill Interamericana de España.

León, O.G. & Montero, I. (2001). Cómo explicar el concepto de interacción sin estadística: Análisis gráfico de todos los casos posibles en un diseño 2x2. *Psicothema*, 13, 165-171.

Moreno, S. & Sánchez, A. (1998). Análisis bibliométrico de la revista "Psicothema" (1989-1997). *Psicothema*, 10, 23-27.

Shaughnessy, J.J., Zeichmeister, E.B. & Zeichmeister, J.S. (2000). *Research methods in Psychology* (5th ed.). New York: McGraw-Hill.

## APPENDIX A

### CATEGORIES USED IN THE CLASSIFICATION SYSTEM

The system of classification of the methodologies used in the different studies analyzed was developed on the basis of the terminology proposed by León and Montero (1997). Below we present the eight main classification categories and their subcategories, where applicable. Given that the denominations are sufficiently well known, we only provide inclusion criteria in cases where, in the literature, the denominations are not common from author to author.

**Theoretical study:** we included in this category all the articles that did not provide empirical data generated by the authors. That is, all reviews that did not report an empirical study.

**Observational descriptions:** This category includes studies that used natural or structured systematic observation, with a descriptive objective.

**Survey study:** This group included all the studies that used surveys with a descriptive purpose, indicating the type of design used, transversal or longitudinal.

**Case study:** We considered as case studies all those dealing with a single sample unit, be it a person, group, organization or other, and whose objective was descriptive.

**Experimental study:** In order for a study to be classified as experimental it was necessary for at least one of the variables studied as independent to have been manipulated by the researcher. Experiments were categorized as simple – a single independent variable – or complex – more than one independent variable. We also noted the nature of each one of the independent variables – manipulated or not manipulated – and the type of design – between- or within-subjects – employed with each one.

**Quasi-experimental study:** In this category we included all those studies that, despite having the objective of checking a hypothesis of causal relation, had limitations – more or less serious – for successfully accomplishing that objective. We included here designs with intervention – applications in natural situations in which it is impossible to assign subjects at random or control the order of application of the levels of the independent variable – and "ex post facto" studies, characterized by the impossibility of manipulating the independent variable. Within the designs with intervention we took into account each of the thirteen variants listed in León and Montero (1997). We made the same consideration for the "ex post facto" studies.

**Single-subject experimental study:** In this category we included all the studies that used existing experimental techniques for applica-

tion to single cases. Five variants were found: AB, ABAB, multiple baseline for various behaviours of the same subject, multiple baseline for the same behaviour in various situations, and multiple baseline in various cases.

**Instrumental study:** We considered as belonging to this category all the studies whose purpose was to develop tests and equipment, including both their design and adaptation.

## APPENDIX B

### REFLECTIONS FOR REVIEWERS

These lines are addressed to reviewers of the journal and not to authors, since it is the former that decide whether a manuscript is ready for publication. We feel that the communicability and replicability of reports would be improved if reviewers were to take into account the following and to inform the authors in a detailed and didactic way. We should stress that the comments that follow refer to exceptions, and do not correspond to either the average or the mode of the publications.

It is obvious that the design serves to achieve an objective made explicit in the introduction – so obvious that we as authors often think it unnecessary to explain it or justify it. The reader tends to have a different point of view.

**Common ground:** we do not like norms. This conclusion would be reached by anyone that reads the set of "methods" sections from these ten years of journals. If those who write and review studies were to follow the APA norms (1994) required by the publication, this paragraph would be unnecessary. Specifically, we feel there would be an improvement if we insisted upon: a) naming the design that is to be used; stating whether the variables will be between- or within- and indicating the corresponding number of levels and their names; b) making explicit how groups were formed – making it clear to authors that not using a biased criterion is not the same as a random assignment; c) avoiding the word "influence" in titles and abstracts when the research is not experimental; d) using the recommended sub-sections in "methods" and not adding others (e.g., statistical procedures); e) dispensing with creativity when making tables: it is easier for the reader to follow a standard format; f) reporting the means and variances of the groups; and h) although it may seem punctilious, writing up the statistical results in a standard way: a different format from each author in the same issue of a journal does not give an elegant result.