Five research principles to overcome the dualism quantitative-qualitative

ROBERTO TRINCHERO

Abstract: The article intends to suggest some reflections about the criteria that should guide the design of empirical research and the evaluation of research plans and reports. The main idea is to demonstrate that educational research is intrinsically multi-method and mixed-method. Qualitative and quantitative moments are present in all the research and the two approaches are closely interrelated. To overcome a "paradigmatic" view that impose the rigid separation between quantitative and qualitative approaches, five research principles are proposed. These principles are related to: a) planning of the research; b) declaration of the theoretical “lenses” by which the researcher see (and interprets) the world; c) designing and documenting of the research plan; d) analysis of the data gathered; e) definition of the logical extension of the research results. The adoption of the illustrated principles is a possible way to promote a new “culture of empirical research” in education, based on openness to different theoretical contributions (also from different disciplines), epistemological approaches and operative strategies. A research that aspires to have a decisive impact on current educational practices must be a research able to define not only “what works”, but also “how works”, “why works” and “where works”, and only the synergy between the qualitative and quantitative approaches can aid the research to pursue successful these objectives.

Riassunto: L’articolo propone alcune riflessioni sui criteri che dovrebbero guidare la progettazione e la valutazione di piani e rapporti di ricerca. L’obiettivo è dimostrare che la ricerca educativa è per sua natura intrinsecamente multi-method e mixed-method. Momenti qualitativi e quantitativi sono presenti in tutte le ricerche ed i due approcci sono tra di loro strettamente interrelati. Allo scopo di superare una visione “paradigmatica” che impone la rigida separazione tra approcci quantitativi e qualitativi, vengono proposti cinque principi ricerca. Tali principi riguardano: a) la pianificazione della ricerca; b) la dichiarazione delle “lenti” teoretiche con cui il ricercatore vede (ed interpreta) il mondo; c) la progettazione e documentazione del percorso di ricerca; d) l’analisi dei dati raccolti; e) la definizione dell’estensione logica dei risultati della ricerca. L’adozione dei principi illustrati costituisce una possibile via per promuovere una nuova “cultura della ricerca empirica” in educazione, basata sull’apertura a differenti contributi teorici (anche provenienti da diverse discipline), approcci epistemologici e strategie operative. Una ricerca che aspira ad avere un impatto rilevante sulle pratiche educative correnti deve essere una ricerca in grado di definire non solo “cosa funziona”, ma anche “come
funziona”, “perché funziona” e “dove funziona”, e solo una stretta sinergia tra gli approcci qualitativi e quantitativi può aiutare la ricerca a perseguire con successo questi obiettivi.

**Keywords:** Methodology of research in education; research design; data analysis; mixed-method research; research evaluation.

**Introduction**

Qualitative and quantitative methods of research are really opposed? Several authors have criticized this paradigmatic separation, from the ’70s to today (see for example Perrone, 1978; Reichardt & Cook, 1979; McNamara, 1979; Eckeberg and Hill, 1980; Bryman, 1988; Howe, 1988; Hammersley, 1992; Pring, 2000; Tashakkori & Teddlie, 2003; Niglas, 2008; Bergman, 2008; Gerring, Thomas, 2011), stating that the differences on the ontological and epistemological level are not justified on the methodological level. When the researcher designs a research, he/she needs to take in account several problems related to the validity/confirmability of his/her study. These problems require answers drawn from either the tradition of quantitative and qualitative research. More, quantitative and qualitative practices are intrinsically strictly interrelated: “qualitative” moments are present in quantitative research and “quantitative” moments are present in qualitative research.

In this article, we propose five research principles useful to avoid serious mistakes imputable to a “paradigmatic” view, that imposes a rigid separation between quantitative and qualitative approaches in educational research. These principles are related to: a) planning of the research; b) declaration of the theoretical “lenses” by which the researcher see (and interprets) the world; c) designing and documenting of the research plan; d) analysis of the data gathered; e) definition of the logical extension of the research results.

**1. Plan your research starting from your research objective not from your research method**

Different research objectives require different approaches of research. This simple assertion is not so obvious. In the tradition derived from the
Methodenstreit, qualitative and quantitative approaches are different “methods” (from the Greek *methodos*, “which indicates the path or way”, composed of *meta*, beyond, and *hodos*, walking), derived from different ontological and epistemological visions. This means that the researchers make a choice “a priori”: if he/she chooses to join to the realist perspective he/she becomes a quantitative researcher, if he/she chooses to join to the constructivist/interpretivist perspective he/she becomes a qualitative researcher. This tradition see researchers as “partisans” deployed in favor of quantitative or qualitative approach, seen as a pair of alternative and incompatible methods.

Adopting a pragmatic approach (“Given a contingent research objective, what is the most appropriate way to build valid/confirmable answers?”), researcher can use indifferently the qualitative or the quantitative perspective (or a mix of these perspectives) depending on the objective to pursue. This is the concept of “research strategy”. A “strategy” is a set of techniques and principles for the application, specifically intended to pursue an objective. In the perspective of research strategy (in opposition to the perspective of the research method), qualitative and quantitative tradition are not alternative and opposite approaches but complementary strategies in order to capture the complexity of the real in relation to the objectives of the study. In the perspective of “research strategy”, the researcher may use (and combine in a creative way) techniques and tools for data gathering and analysis typical of the classical quantitative and qualitative approach to maximize validity/confirmability of the study.

Let’s see an example. If the research objective refers to establish the existence of a causal relationship between factors (as in the hypothesis “Motivation to learn can improve scholastic achievement”), the applicable epistemology should refer to the tradition of experimental research with randomized samples. However, this is not enough. Quantitative experimental design can be useful to quantify the increase of scholastic achievement probably due to a corresponding increase of motivation, but how can we demonstrate that increase of scholastic achievement is really due to increase of motivation and is not imputable to other extraneous factors? The only way is to study the dynamics that led an increase of motivation to promote an increase of achievement. This is possible by a persistent engagement of the researcher in a qualitative in-depth inquiry on the field in which the experiment was performed, to reveal the processes underlying the causal relationship and the conditions under which this relationship may occur. Experimental design can help us to say if an intervention work
and in *which measure*. In-depth observation can help us to say *what* really work, *why* and *where*. Thus, a good experimental research must use techniques drawn from both quantitative and qualitative approaches.

Figure 1 show some examples of research strategies that combine qualitative and quantitative approach to respond effectively to the listed research objectives.

**Fig. 1 – Examples of research objectives and possible related research strategies**

<table>
<thead>
<tr>
<th>Research objectives</th>
<th>Possible strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Describe a phenomenon and its evolution by means of quantitative and/or qualitative parameters</td>
<td>Quantitative inquiry (Hyman, 1955), for quantitative surface description. Interpretive research (e.g. Qualitative inquiry, Phenomenological inquiry, Ethnography, Biographical approach, etc., see Denzin, Lincoln, 2002), for qualitative in-depth description. Case study research (Stake, 1975), for qualitative/quantitative in-depth and evolutionary description.</td>
</tr>
<tr>
<td>Explain a phenomenon on the basis of other phenomena / Predict a phenomenon from a set of initial conditions</td>
<td>Quantitative inquiry, for correlational studies. Experimental research (Campbell, Stanley, 1963), for causation studies, complemented by interpretive research, to reconstruct the dynamics at the base of the causation. Grounded theory (Glaser, Strauss, 1967), for building theoretical frameworks that explain a phenomenon.</td>
</tr>
<tr>
<td>Understand the &quot;good reasons&quot; that stands behind actions, choices and human attitudes / Predict actions, choices and human attitudes from a set of initial conditions</td>
<td>Interpretive research, to understand phenomena through accessing the meanings that participants assign to them.</td>
</tr>
<tr>
<td>Interpret a phenomenon in the light of a given theoretical framework or other phenomena</td>
<td>Interpretive research, to reconstruct meanings associated to the phenomenon on the basis of a given theoretical framework or set of phenomena.</td>
</tr>
<tr>
<td>Compare more phenomena, in the same or in different context</td>
<td>Comparative inquiry (Ragin, 1987), to establish what is common and what is different in a set of phenomena.</td>
</tr>
<tr>
<td>Identify possible solutions to contingent problems encountered in a given context</td>
<td>Action research (Kemmis, McTaggart, 1982), to construct &quot;good&quot; solution to problems that arise from field practice.</td>
</tr>
</tbody>
</table>
2. Make explicit the “lenses” with which you look at and interpret the reality

Is possible to observe a phenomenon leaving aside our preconceptions and “models of the world”? Objectively not. When we observe and assign meaning to a phenomenon we use a range of interpretative models, more or less explicit. Each observation contains theory and models of a theory (Gallino, 1987). If our observation, description and classification are not driven by explicit theories and models, the risk is that these are driven by several uncontrolled implicit models. The consequence is to gather incomparable data, poorly interpretable and combinable with each other to build a coherent conceptual framework. The data are not “given” but are the outcome of a process of construction that reflects the mental models of the observer. In this sense the data are “theory-laden”, according to the famous expression of Kuhn (1962). The observer gives definitions for the concepts, identifies indicators, builds and applies data gathering tools, selects what elements are significant for the research and what are not.

To make “controllable” the research process, qualitative and quantitative researchers must declare theoretical framework for the study, working hypotheses, factors and concepts considered important for the study, and their conceptual and operational definitions. These are the “lenses” with which the researcher looks at the reality. In a research about the mathematical competence of the students, what is the conceptual definition that the researcher gives to the term “mathematical competence”? What are the indicators useful to observe mathematical competence in students? What are the expectations that the researcher has about the mathematical competence of students (hypotheses)? A good operational definition make explicit “what to observe”, “how to observe” and “how to interpret what is observed”.

Studies that share the same theoretical framework and give to concepts the same definitions are mutually comparable and can produce cumulative knowledge. Research results produced by studies that adopt different theoretical frameworks, concepts and operational definitions are not directly comparable because are derived from different models of selection and assignment of meaning to phenomena. Either qualitative and quantitative approaches must deal with the problem of meaning of concepts involved in the study and the problem of selection of what is significant to study a reality. In this process, classical qualitative and quantitative researchers proceed
in different but related ways. Quantitative researchers tend to maximize comparability of the observations (see Gerring, Thomas 2011), because quantitative epistemology use subjects of the study as witnesses of a relation between factors and so the quantitative data analysis require a single-dimensional scale and a precise and exhaustive categorization of each observation. Qualitative researchers don’t need to assume a high level of comparability among observations because qualitative epistemology insist on the specificity of the subjects of the study, that also emerge from the differences and commonalities between them. For the quantitative researcher the variety of meanings that the words can assume is a serious problem (e.g. in a classical Likert satisfaction scale: 1 Very Satisfied – 2 Satisfied – 3 Neutral – 4 Unsatisfied – 5 Very Unsatisfied, if two subjects use a different meaning for the word “Satisfied”, their answers are obviously incomparable). For the qualitative researcher the variety of meanings that the words can assume for different subjects can be instead a formidable research opportunity and, if his/her research objective is the in-depth description, his/her efforts are not concentrated on comparability but on the specificity of single subject. Differing assumptions about comparability may be grounded in differing ontological and epistemological views, but comparability is not inherent to the nature of things: only “lenses” of the researcher can see the need/opportunity of a “quantification” or a “qualification” and only his/her conceptual operations can produce them. Qualitative and quantitative data gathering lies on a continuum, depicted in Figure 2.

In qualitative research, if the study insists on the same properties for all the subjects under study, data gathering can originate a comparison grid (semi-structured data). For each cell in the grid, a specific text describes the form that the related property assumes for the subject considered (e.g. cultural traits of the subject, attitudes, choices, behavior in determinate situations, etc.). If the considered properties are different for the various subjects (typical for the in-depth description), data gathering originates long textual description (low-structured data). In both cases, the ways in which human beings make sense of their lived experience can be extraordinarily diverse through time, across cultures and individuals. The use of textual description avoids to reduce this complexity to standard categories and facilitates in-depth analysis.

In quantitative research, the same concepts and properties are operationalized through an operative definition, made comparable through classification into pre-defined categories originating variables in a data matrix.
Five research principles to overcome the dualism quantitative-qualitative

Fig. 2 – A possible continuum between data matrix and textual in-depth description

<table>
<thead>
<tr>
<th>High-structured data</th>
<th>Semi-structured data</th>
<th>Low-structured data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data matrix</strong></td>
<td><strong>Grid of comparison</strong></td>
<td><strong>Textual in-depth description</strong></td>
</tr>
<tr>
<td>Variable 1</td>
<td>...</td>
<td>Subject 1: Text...</td>
</tr>
<tr>
<td>Variable 2</td>
<td>...</td>
<td>(Property 1, 3, 7, 8)</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>Subject 2: Text...</td>
</tr>
<tr>
<td>Case 1</td>
<td>Subject 1</td>
<td>(Property 2, 4, 5)</td>
</tr>
<tr>
<td>Data...</td>
<td>Text...</td>
<td>Subject 3: Text...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Property 1, 4, 6, 9, 10)</td>
</tr>
<tr>
<td>Case 2</td>
<td>Subject 2</td>
<td>...</td>
</tr>
<tr>
<td>Case 3</td>
<td>Subject 3</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td></td>
</tr>
</tbody>
</table>

For each cell in the matrix, a number (e.g., “20”) or a category label (e.g., “Male”) resumes with a single value the state that the related subject assumes on the relevant property, and this state assumes a mean in relation to the empirical phenomenon that it represents (e.g. the number “20” is not a quantitative observation, the statement “John is 20 years old” is a quantitative observation, because “20” is a precise state on the property “age” corresponding to subject “John”). The presence of an operational definition forces the observer or the respondent to a questionnaire to make explicit choices in order to classify phenomena in pre-defined categories.

In conclusion, qualitative and quantitative observations are inevitably rooted in language and strictly depend from the conceptual and operative definitions of the key-terms adopted in the study, which in turn depend on...
the theoretical framework adopted, that in this manner influence significantly the course of the study. In addition, meanings that subjects give to words are an important element to be considered both in qualitative and in quantitative studies.

3. Describe in-depth your research process and construct appropriate counterargumentations to make him stronger

The researcher should describe and justify the research logic, referring to the description of investigations in all its phases, from the question that originates the study to the conclusion of the research report. This argumentation must be explicit, logically consistent and resistant to counterarguments. What are the possible “bias” (e.g. intervening factors not considered in the study, non-randomization of the samples, problems of non-response, errors in the research design, data gathering, analysis and interpretation) of the research results? A rigorous counterargumentation process performed by one or more peer reviewer can surely improve the “solidity” of the statements generated by research and broadening their scope. But in this validation process the first “counterargumenter” should be the researcher itself. Designing research having in mind possible source of criticism (and declare them in the discussion) helps to carry out a more conscious and a more solid work. Figure 3 shows some examples of questions useful to formulate counterarguments for a research report of a quantitative inquiry. The same guidelines can be used to assess and improve a research project in the design phase.

Fig. 3 – An example of a guide to the counterargumentation process for a descriptive or correlational study

1. Questions, topic, objectives. The research report describes the knowledge needs that originate the research (research questions) and the intent of the researcher (research objectives)? Defines the research topic? Contextualizes questions and objectives into a broader line of studies?

2. Theoretical framework. The research report describes the theoretical framework of the research, illustrating authors and theories that support the research decisions? The theoretical framework is consistent with the research topic and the declared research questions and objectives?

3. Working hypothesis. The research report make explicit the working hypothesis that have guided the research? These assumptions are formulated as answers to the research questions? These answers are consistent with the theoretical framework?
Five research principles to overcome the dualism quantitative-qualitative

4. **Factors and relations between them.** The research report explicit the factors taken into account in the study? These factors are effectively formulated in the same manner in which they are formulated in the working hypothesis? Is made explicit which factors are dependent, independent, moderator, intervening? Is made explicit the type of relationship between independent and dependent factors (e.g., covariation, causation, etc.)? The study describes the way in which the possible intervening factors may act on the dependent factors and the way in which the possible moderator factors can influence the relationship between independent and dependent factors?

5. **Operational definition.** The research report provides an operational definition in which is made explicit, using the theoretical framework, the correspondence between factors and indicators? Is made explicit the correspondence between indicators and items in the instruments used to gather the data (e.g., the correspondence between indicators and items in the questionnaire)? Is made explicit the correspondence between items and variables in the data matrix? Are made explicit the rules adopted for the construction of the indices that summarize the data? The indicators used in operational definition are specific, observable and measurable properties of the subjects under study (e.g., personal characteristics, behaviors, choices, attitudes, knowledge, skills)? Are they defined in precise, unambiguous terms that describe clearly and exactly what is being observed?

6. **Population and sample.** The research report describes the characteristics of the population taken in account in the study? Are described the characteristics of the sample (numbers of subjects, sampling technique, possible bias in the selection) on which the study was conducted? Are described the procedures used to integrate the sample in the case of rejection of the subjects to participate to the study?

7. **Techniques and instruments to gather the data.** The research report describes the techniques (e.g. interviews, session of observation) of data gathering adopted and the instruments (e.g. questionnaire, check list, rating scales) used to apply them? The choice of these techniques and instrument is correctly justified? There is a specific reflection about the possible sources of invalidity and unreliability related to techniques and instruments used?

8. **Data gathering process.** The research report describes in detail the procedures followed to establish contact with the subjects of the sample and to collect the data for the study using the techniques and instruments described above? Are made explicit the eventual actions of triangulation used to increase the reliability of the study? Are made explicit the possible sources of invalidity and unreliability associated with the process of data gathering?

9. **Data analysis.** The research report make explicit the techniques used for the data analysis? The choice of such techniques is justified in relation to the research objectives and to the available data? The techniques used are adequate to the levels of measurement of the variables? The significance of the results is considered? Are made explicit and justified the logical steps between the results of the data analysis, the testing of the working hypotheses and the conclusions of the study? Alternative interpretations of the results of the study are proposed?

10. **Discussion and meta-reflection.** The research report presents a paragraph of discussion of the results obtained in the study? In the discussion, it takes into account alternative explanations for the result obtained? It defines the limits of extensibility of the results? It makes explicit the knowledge gained by the researcher due to study conduction (if the researcher would redo the research, what will redo in the same way and what will redo in a different way)? Are made explicit the strengths and weaknesses of the study, also in relation to similar studies? The research report offers guidelines for further studies with similar research question and objectives?

In this counterargumentation process, the researcher must operate an in-depth description of the “good reasons” that underpin the decisions tak-
en by the researcher itself. This reflective activity has clearly many common points with interpretive research. The conclusion is: also the quantitative research requires a self-reflection process that employs conceptual frameworks, techniques and instruments typical of qualitative research.

4. Analyze your data using both quantitative and qualitative frameworks

Data analysis always requires both a quantitative and a qualitative sensibility. The interpretation of quantitative data present several “traps” and the researcher must be aware of them (see Marradi, 1988; Marradi, Gasperoni, 1992 and 2002; Miceli, 2004). In the following paragraphs we present some common traps in data analysis and a proposal of a technique that help us to analyze “quantitative” data with a “qualitative” approach.

4.1. Be aware of the correlation and causation traps

The correlation coefficient is a basic statistic for many bivariate and multivariate techniques of data analysis that involve cardinal data (e.g. factor analysis, path analysis, etc.). But how reliable is the number represented by this coefficient? Consider the following example illustrated in McMillan and Schumacher (2010, 234): in the scatterplot graph labelled by A (see Figure 4) the presence of an outlier (an “aberrant” value, a value which deviates abnormally from the group) increases the coefficient of correlation between the two variables to the value r = 0.67. The researcher that formulate an interpretation based only on the value of this coefficient can conclude that there is a very high correlation between the two variables. But examining the scatterplot it is very easy to see that this interpretation is erroneous, because this value of correlation coefficient is due to outlier. By removing that person from the sample the correlation coefficient decreases to r = 0.09, and this lead us to the correct interpretation: there is no correlation between the two variables. This removal requires a justification necessarily based on arguments constructed with interpretive research framework: what are the “good reasons” for which the subject is not comparable with the other subjects of the sample? Graph B of Figure 4 depicts the opposite situation: an accurate examination of the scatterplot shows a
clear relationship between the two variables, summarized from the value of the correlation coefficient ($r = 0.84$). By placing into the sample a single outlier, this value drops to $r = 0.23$. Also in this second case, if the analysis is limited to the “quantitative” correlation coefficient, the researcher may run into a serious error. Only with an accurate reasoning about the “qualitative” aspects of the subjects in the distribution, the researcher can uncover the “good reasons” that lead one or more subjects to deviate from the “trend” expressed by the entire group.

*Fig. 4 – The correlation coefficient trap (elaborated from McMillan, Schumacher, 2010, 234)*

Let’s see another trap. If the situation is as illustrated in the graph B without outlier, we can say that there is a high correlation between the two variables in mathematical sense. But can we say that there is a strong association between the two factors or phenomena quantified by the two variables? We can’t get this conclusion if we are reasoning only on the “mathematical” representation of phenomena. For example, if the variable in abscissa is an index of presence of storks in a territory and the variable in ordinate is an index of human births in the same territory, can we say that there is an association between presence of storks and rate of birth
in a territory? Obviously not, because our “qualitative” knowledge tell us that there is no logically necessary connection between presence of storks and rate of birth. Only our “qualitative” knowledge of the situation under inquiry suggest us that there is a third variable (the rurality of the area) that is logically associated with both factors and can explain this paradox (Lazarsfeld, 1955).

In addition, if the logic argumentation demonstrates that there is an association between two factors or phenomena, this doesn’t mean that the first factor of phenomenon cause the second or vice-versa. For example, the statement “the improving of results in mathematics of a group of students is due to introduction of a new teaching method” can be considered corroborated by data if: a) there is a mathematical correlation between the index of mathematical results and the index of application of new teaching method (and these indexes are derived from good indicators of the two phenomena); b) there is a defensible logical connection between the two phenomena, that let us to demonstrate that changes in one leads to logically related changes in other; c) the phenomenon supposed as cause of the other is happened before the phenomenon supposed as effect; d) it is possible to demonstrate that the changes in dependent phenomena are not related to changes in other phenomena logically correlated with this. For these reasons, the demonstration of causal relationship requires experimental design on randomized samples with a “quantitative” analysis to demonstrate mathematical correlation and a “qualitative” analysis to argument logical connection between factor under consideration and to exclude that changes in dependent factor are imputable to other phenomena.

4.2. Be aware of the X square trap

The correlation trap refers to cardinal data analysis. Analog considerations can be done for categorical data analysis. To establish the existence and quantify the strength of the relationship between two categorical variables, we traditionally use several indices of association between variables (e.g. phi, phi square, Cramer’s V) derived from the statistical index $X^2$, based on the difference between observed and expected frequencies in each cell of the contingency table generated by joint frequency distribution of the two variables. More this index is away from zero, more the relationship between two categorical variables is considered signifi-
cant. In the table A of Figure 5, we have a contingency table that classifies 57 subjects, male and female, in three level of outcome in a language test (Basic, Intermediate, Advanced). In the table B are illustrated the expected frequencies for each cell. The value of X square for the table is 10.33 with a significance of 0.0057. The apparent conclusion is “in the sample, males and females have obtained significantly (at 0.01 level) different levels of outcome”. However, if we examine the standardized residuals\(^2\) for each cell (table C, Figure 5), we see that differences between observed and expected frequency are not significant, except for the cell “Female-Basic Level” that has a standardized residual of -2.0 (significant at 0.05 level). So, the apparent conclusion is wrong. Data tell us that the high value of X square is due to the fact that we expect 7.5 female with Basic level of outcome and we found only 2 cases. The correct conclusion is “in the sample, females tend to not leave the course with a Basic level of outcome”, significant at the 0.05 level. Nothing of statistically significant may be stated about the gender of the students with Intermediate and Advanced level.

Also for categorical data analysis, a correct interpretation should not stop at the superficial reading of the “quantitative” index, but requires a “qualitative”, in-depth, reflection about the data distribution in relation to the theoretical model under control (in this case the model of the statistical independence of the values assumed by the two variables).

**Fig. 5 – The X square trap**

<table>
<thead>
<tr>
<th>Table A: Observed frequencies</th>
<th>Table B: Expected frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>23</td>
</tr>
<tr>
<td>Female</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>25</td>
</tr>
</tbody>
</table>

\[
\text{X square } = 10.33 \quad \text{sign. } = 0.0057
\]

<table>
<thead>
<tr>
<th>Table C: Standardized residuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lev. B</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Female</td>
</tr>
</tbody>
</table>

EDUCATION SCIENCES & SOCIETY
4.3. Consider alternative to classical data analysis: the Rough Data Set Rule Extraction procedure

A correct application of classical techniques of data analysis requires both “quantitative” and “qualitative” sensibility. There are also techniques of data analysis which are situated on the border between the “quantitative” and the “qualitative” side. An example is the data mining technique called Rough Data Set Rule Extraction (Pawlak, 1991). This technique allows, starting from a common matrix cases by variables, the extraction of patterns, that are sequences of values which are repeated regularly in the subjects present in the matrix. In the example of Figure 6, to a large sample of adolescents were asked to express their agreement or disagreement with the statements expressed by characters within 5 short “vignettes” relevant to the normal teacher-student interaction, narrated in a questionnaire. Agree/disagree with these statements are indicators of the attitudes of the students toward the teachers in relation to 5 involved topics. Combinations of these attitudes define specific “profiles” of attitudes. The application of the Rough Data Set Rule Extraction procedure has generated three profiles with percentage weight greater of 5% (all others profile are considered marginal profiles) of the sample, as illustrated in Figure 6. These profiles represent a synthesis of the views expressed by students and can be used in a descriptive perspective (if the research objective is the description of the opinion profiles of the sample) or in a comparative perspective (if the research objective is to describe commonality and differences between opinion of subgroups, e.g. males and females). It is also possible to control specific hypotheses by the quantification of the percentage of subjects whose opinion is consistent with the hypothesis under examination.

When we analyze data with the Rough Data Set Rule Extraction procedure, we are doing a quantitative or a qualitative research? We do both research at the same time: “quantitative” because we use statistical procedures and indexes, “qualitative” because we use the principles of the logical inferences and considerations of similarity/difference between data that represent “qualitative” positions of the students.
Five research principles to overcome the dualism quantitative-qualitative

In summary, “qualitative” and “quantitative” moments are necessarily strictly intertwined in common data analysis. A good research process must take in account both frameworks. In data analysis, quantitative and qualitative are “mixed-method” by their nature.

5. Define the logical extension of your research results by grasping how, why and where they are valid and confirmable

A research result is statistically generalizable only if the sample can be considered representative of the population from which it was extracted, but is it possible to construct a really representative sample? Marradi (1997, 25) states that we can establish the representativeness of a sample only when we have the necessary information related to the population (in which case the corresponding information on the sample is a duplicate). If we do not have the relevant information about the population (and if we had we would not need to do research…), we can’t say anything about the representativeness of the sample. This is the so-called “paradox of representativeness”. The random selection of the samples under study (randomization) can ensure us that there are no systematic biases (due to

---

**Fig. 6 – Profiles of attitudes towards the teachers obtained by the Rough Data Set Rule Extraction procedure**

<table>
<thead>
<tr>
<th>Profile</th>
<th>Weight</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teachers <em>can</em> help students to deal with their discomfort</td>
<td>256</td>
<td>23.5</td>
</tr>
<tr>
<td>If students do not pay attention in class the fault <em>is not</em> of the teacher</td>
<td></td>
<td></td>
</tr>
<tr>
<td>If the teacher do a mistake this is not a serious problem</td>
<td></td>
<td></td>
</tr>
<tr>
<td>If the teacher constantly criticizes a student the fault is of the student</td>
<td></td>
<td></td>
</tr>
<tr>
<td>If a student gets systematically bad grades should try to talk to the teacher</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teachers <em>can’t</em> help students to deal with their discomfort</td>
<td>157</td>
<td>14.4</td>
</tr>
<tr>
<td>If students do not pay attention in class the fault <em>is</em> of the teacher</td>
<td></td>
<td></td>
</tr>
<tr>
<td>If the teacher do a mistake this is not a serious problem</td>
<td></td>
<td></td>
</tr>
<tr>
<td>If the teacher constantly criticizes a student the fault is of the student</td>
<td></td>
<td></td>
</tr>
<tr>
<td>If a student gets systematically bad grades should try to talk to the teacher</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teachers <em>can’t</em> help students to deal with their discomfort</td>
<td>57</td>
<td>5.2</td>
</tr>
<tr>
<td>If students do not pay attention in class the fault <em>is</em> of the teacher</td>
<td></td>
<td></td>
</tr>
<tr>
<td>If the teacher do a mistake this is not a serious problem</td>
<td></td>
<td></td>
</tr>
<tr>
<td>If the teacher constantly criticizes a student the fault is of the student</td>
<td></td>
<td></td>
</tr>
<tr>
<td>If a student gets systematically bad grades should try to talk to the teacher</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
the sampling process) in the sample selection and gives us the possibility
to calculate the margin of statistical error (confidence intervals) in the es-
timation of population parameters from the sample parameters, but does
not guarantee neither the representativeness of the sample nor the perfect
equivalence of the sample with other samples drawn from the same popu-
lation. In addition, the populations are entities in continuous evolution.
Even if it were possible to extract a truly representative sample from the
population of Italian teenagers, who assures us that the extracted sample
in 2011 is still representative of the teenagers of 2012? To generalize a re-
search result obtained on a sample to a not well-defined "population" seems
to be an ambition rather than a real possibility.

So, how it is possible to extend a research result obtained on a sample
to other samples or to entire population? A more reliable way might be
the definition of the logical extension of the research results, that is the set
of conditions that must be present because it can be observed in the new
sample (or in a broader population) the same associations between factors/
phenomena observed in the study sample. For example, we can observe
some effects on mathematical ability in a sample of children who have
attended a basic chess course and we can demonstrate some correlation
between improve on mathematical ability and hours of frequency in chess
course. To extend this result to other samples we must to define the condi-
tions under which this correlation can have place. This set of condition is
the logical extension of the result obtained.

To define the logical extension of a results is necessary a “thick” de-
scription of: a) the mechanism that, according to the researcher, causes the
association between factors/phenomena; b) the conditions under which
the result has been obtained; c) the essential conditions, in the absence
of which, the association between factors/phenomena can’t be observed.
For the chess and mathematics research example, the supposed mechanism
that cause the association between improve of mathematical ability and
hours of the course is an increase of metacognition ability (due to practice
in chess game and to the approach of the instructor) that ameliorates the
mathematical problem solving ability of the children. The condition un-
der which the results has been obtained refers to in-depth description of
the research context and process (diary of experiment, participant observa-
tion in class during the chess experience, etc.), and the essential conditions
refers to various research experience in which some parameters can vary
(e.g. duration of the chess course, approach of the chess instructor, sup-
port given by class teacher, computer-based chess practice, initial condition of the children, etc.) that consent to state that to observe some effect on mathematical ability the chess course must have a duration at least of 16 hours, the chess instructor must use a precise protocol, children must practice their skills playing with a special software reaching at least the level 7 of game, etc. In presence of these conditions, the research results can be applicable to other contexts.

Obviously, in the human sciences, the full control (or the simply take into account) of all the possible factors that may affect a result is impossible, and also impossible is the exact replication of the study in other contexts, simply because each research context has his specificity. Adopting the described approach, the factors/phenomena classically called “threats to validity of the study” are instead opportunities to extend the validity/confirmability of the study, while the replications of the study in different contexts are opportunities to extend or reshape the set of conditions that defines both the limits of validity/confirmability of the study and the logical extension of research results. The aim is to distinguish the “factors/phenomena that may affect outcomes” from the “factors/phenomena that can’t affect outcomes” in order to define the conditions of validity/confirmability of the study and so the logical extension of the results.

According to this approach, the research questions ought to be not only “There is an association between the two considered factors/phenomena?” or “How much strong is this association?”, but “What are the dynamics underlying this association?” “How this association works?”, “Why there is this association?”, “Where this association can have place?”. Finding answers to these questions needs both the classical approach of the “quantitative” correlational studies and the approach of “thick”, in-depth description of the dynamics underlying the phenomena, strongly inspired to the classical “qualitative” studies. Quantitative approach can help us to define “what works” and “what have the higher effect”, but only qualitative approach can help us to define “how works”, “why works”, “where works” and so help us to transfer and apply a research finding to other contexts.

6. Non-Conclusion

Doing research is primarily a reflective activity, which can’t be neither reduced to the mere application of procedures nor evaluated solely on the
basis of the presence or absence of certain requirements. “Anchor to the data to go beyond the data”, should be the watchword of every researcher. “Going beyond the data” means to see what is not immediately visible, perform a creative act that reorganizes the existing conceptual framework in a new and unexpected manner, devise new solutions to old problems. A good empirical research, powered by methodological rigor and creative tension, is a research that is really able to change the current practices. But to change the practices, is necessary to know them in depth, not simply at the surface level. This requires from the researcher an attitude of openness to different theoretical contributions (e.g. contributions from different disciplines), epistemological approaches and operative strategies. In this view, the research is intrinsically multi-method and mixed-method, because qualitative and quantitative moments are strictly interrelated in real research context. The five principles proposed are intended to be a stimulus to the debate and a way to entice new researchers to build new methods that let us to go beyond the existing. In this sense, this reflection is not intended to be an end but a beginning: not a conclusion but a non-conclusion.

Author’s Presentation: Roberto Trinchero is Associate Professor of Experimental Pedagogy and Methodology of Educational Research at the Department of Philosophy and Educational Sciences of the University of Turin. His research interests include the methodology of research in education and the design and evaluation of training programs. He is member of the board of the SIPED – Italian Society of Pedagogy.

Notes

1 “X square” is the abscissa value of the chi-square probability distribution.
2 A Standardized residual (or Studentized residual) is the quotient resulting from the division of a residual by an estimate of its standard deviation. In this case is the difference between observed and expected frequency divided by the square root of the expected frequency: (Observed frequency – Expected frequency)/sqrt (Expected frequency). This value is significant at 0.05 level when his absolute value is greater than 1.96.
3 This technique is implemented on the package JsStat (www.edurete.org/js-stat).
4 See research reports in Trinchero (2012 and 2013).
References


— (2002), Educational research: Planning, conducting, and evaluating quantitative and qualitative research, Upper Saddle River, Merrill Prentice Hall.


WHEN EDUCATION RESEARCH MEET OTHER TOPICS

Kemmis, S., McTaggart, R. (1982), The action research planner, Geelong, Victoria, Deakin University Press.

Kuhn, T. (1962), The Structure of Scientific Revolutions, Chicago University Press.


Five research principles to overcome the dualism quantitative-qualitative

