

ORIGINAL ARTICLE

## *Exploring challenging classroom behaviors from teachers' perspective: a mixed-method study*

Stefano Castelli<sup>1-A,C,D,E</sup>, Loredana Addimando<sup>2-B,E,F</sup>, Alessandro Pepe<sup>1-B,C,D</sup>

1: Università degli Studi di Milano-Bicocca, Milano, Italy

2: La Scuola universitaria professionale della Svizzera italiana (SUPSI), Locarno, Switzerland

### BACKGROUND

In order to explore the representation of challenging students' behaviors through teachers' words, the present paper reports Italian teachers' replies to the open-ended instruction "Describe the behavior of the student you find most challenging in your class", analyzed through a mixed-method approach.

### PARTICIPANTS AND PROCEDURE

The respondents were teachers from primary and lower secondary schools ( $N = 518$ ) in the city of Milan and in other urban and sub-urban areas of the Lombardy Region, Italy.

### RESULTS

Results from correspondence analysis of the open question suggested two main factors. The first factor, labeled perceived willfulness, seems to confirm that in the appraisal of challenging behaviors teachers' personal beliefs about the origin of students' misconducts play an important role. When behaviors are perceived as "involuntary" (such

as learning difficulties), teachers seems to play a more care-taking role in comparison to conducts that are viewed as voluntarily destructive. The second factor can be described as the impact that behaviors have on the teaching process.

### CONCLUSIONS

The qualitative analysis seems to accurately capture even the nuances of what is portrayed by the quantitative survey and helps us in understanding why, in other Italian studies, the Weak student emerged as the most difficult student for Italian teachers: if this population finds it difficult to discriminate between aggressiveness, hostility and hyperactivity, then it is not surprising that this last category gains salience and, maybe, over-representation in surveys.

### KEY WORDS

teachers' stress; word correspondence analysis; mixed-method; teacher-student relationship

---

CORRESPONDING AUTHOR — Prof. Stefano Castelli, Department of Psychology, Università degli Studi di Milano-Bicocca, 1 Piazza dell'Ateneo Nuovo, I 20126 Milano, Italy, e-mail: stefano.castelli@unimib.it

AUTHORS' CONTRIBUTION — A: Study design · B: Data collection · C: Statistical analysis · D: Data interpretation · E: Manuscript preparation · F: Literature search · G: Funds collection

TO CITE THIS ARTICLE — Castelli, S., Addimando, L., & Pepe, A. (2015). Exploring challenging classroom behaviors from teachers' perspective: a mixed-method study. *Current Issues in Personality Psychology*, 3(1), 12–24.

RECEIVED 04.11.2014 · REVIEWED 14.12.2015 · ACCEPTED 19.02.2015 · PUBLISHED 12.03.2015

## BACKGROUND

In the investigation of teachers' perceptions of classroom misbehaviors, a very popular adopted research method has been the survey (Beaman, Wheldall, & Kemp, 2007). Because of its proven reliability and feasibility for implementation on a wide scale, the method is often applied in educational settings. The majority of surveys are based on a very simple schema: a sample of participants answer a pre-defined set of structured or semi-structured materials allowing the researchers to describe the population under study according to one or more theory-based parameters (Dalenius, 1985). Recent developments in the field of educational research (Grasser, 2009) have fulfilled the need to study human behaviors by combining multiple different sources of data (i.e. mixed-method approaches), allowing researchers to integrate an "external" data set with psychometric measurement models. For example, an attempt to quantify qualitative data through the systematic segmentation of subjective verbal production or to qualify quantitative data by collecting discursive materials in tandem with quantitative measures represents a relevant improvement to the tradition of the educational research mainstream. The combination of different approaches (i.e. quantitative and qualitative methods) is useful in revealing multiple pieces of evidence serving as "building blocks" in the research endeavor (Lieberson, 1992).

In order to explore the relationship between students' challenging behaviors and occupational stress, we administered the *Challenging Student Standard Questionnaire* (van der Wolf & Everaert, 2003) along with a set of four open-ended questions to a sample of in-service primary and lower secondary teachers ( $N = 518$ ) with the aim of analyzing data by a mixed-method strategy.

The main rationale is that in answering an open-ended question teachers expressed their first-hand impressions which, under some points of view, can be considered as naïve, common sense opinion. However, we are also prone to think that an "experts' perspective" offers something different than a common view opinion; after all, the Challenging Student Standard Questionnaire was developed in order to gain a *deeper insight* into the phenomenon of teachers' stress and students' challenging behaviors. And a *deeper insight* is not always in accordance with a *common sense* view. It must be remembered that teachers' "diagnosis" (Tillema, 1995) of causes related to classroom behaviors is crucial, because teachers play a primary role in identifying such problems (Zimmerman, Khoury, Vega, Gil, & Warheit, 1995).

We should further take into consideration that an open question implies a recollection task, which is very different from the recognition task associated with a closed question: two radically different

psychological mechanisms (Lorenzi-Cioldi, 1996; Schuman & Presser, 1981). A famous example illustrating this issue is presented by Schuman and Presser (1981), who provide some data regarding the educational values favored by parents for their children: when a closed format was presented, 60% of parents chose the item "independent thinking", but only 5% mentioned the same answer when "measured" with an open format.

Open questions (and verbal transcriptions in general) are proven to be more useful in exploring the salience of a certain theme. At the same time, quantitative assessment is useful when one is interested in exploring "how often" or "to what extent" an empirical phenomenon, or its impacts on some aspects of the respondents' life, was perceived.

Since we are interested in exploring the common-sense representation of the *challenging student* as reported by teachers, in the present study the instruction "Describe the behavior of the student you find most challenging in your class" was taken into consideration and analyzed through word correspondence and cluster analysis in such a way that psychometric measures of the Challenging Student Standard Questionnaire provided the empirical base to group respondents according to measures of the most challenging behavior they deal with.

## THEORETICAL BACKGROUND

### STUDENTS' CHALLENGING BEHAVIORS

Research in teachers' perceptions of undesirable classroom behaviors is a traditional topic in educational psychology (Langfeldt, 1992) because pupils' misbehaviors often present a major concern to teachers attempting to provide effective learning (Soodak & Podell, 1994).

In a very general fashion, a *classroom misbehavior* was defined as any type of behavior that threatens the flow of academic performance (Turnuklu & Galton, 2001). An alternative conceptualization is offered by Emerson (1995), referring to "a culturally abnormal behavior(s) of such intensity, frequency or duration that the physical safety of the person or others is likely to be placed in serious risk" (p. 4). Finally, Burden (1995) included in this category of behaviors any student behavior that is perceived by the teacher to compete with or threaten the academic actions at a particular moment.

Perhaps the most interesting aspect of all proposed definitions is the idea that a *challenging behavior* cannot be considered ontologically problematic but it starts to affect the workflow of learning and impact on teachers' work when it is *perceived* as something negative and/or destructive according to the cultural and social background in which teachers

and students interact. From this point of view, different social norms dominating in each culture (regardless of macro, meso or micro levels of analysis) contribute to shape teachers' reactions in such a way that the same behavior may be seen as more or less desirable and more or less manageable (Weisz, Weiss, Han, Granger, & Morton, 1995).

The adoption of such a theoretical framework leads to three different and interrelated considerations. First, researchers' focus should be shifted from the intrinsic meaning of the challenging behavior itself to the potential meaning that such behavior has in a particular cultural context. As Weisz (1989) has already argued, adults' opinions on children's behavioral problems are influenced by the cultural milieu in many different ways. For example, societal values and personal expectations about children's most appropriate conduct affect adult judgments as to whether a particular child has shown particular problems. The main consequence of such a perspective is that a particular behavior can be perceived problematic in one setting but not in another (Leung & Ho, 2001) and that a behavior becomes *problematic* when it is *troublesome* to someone (Jones, Charlton, & Wilkin, 1995).

Second, from a socio-constructive approach, researchers' attention should be focused on a teacher-student relationship that contributes, step by step, to development of a "behavioral discourse" involving both actors during their daily interactions in the classroom. Pupils were highly observant of the representations of their teachers: they seem to exactly replicate the sort of conducts that are expected of them (Diamond, 1991). In such "discourses", a very important role is played by reciprocal causal attribution processes performed by subjects concerning the sources of challenging behaviors. For instance, a relevant study (Miller, Ferguson, & Byrne, 2000) about pupils' causal attribution of their own misbehaviors in a sample of 11- and 12-year-old students revealed that students offered four different explanations of their conducts: fairness of teachers' actions, pupil vulnerability, adverse family circumstances and strictness of classroom regime. In contrast, a teacher-student relationship can be described as "positive" when language teachers are viewed as tolerant and mathematics teachers are less authoritarian but cooperative (Van Petegen, Aelterman, Van Keer, & Rosseel, 2008). From this viewpoint, teacher-student relationships as well as teachers' perception of challenging behaviors fit within attribution theories dealing with how people perceive the causality of the events and how they explain the origins of a particular event (Weiner, 1992).

The adoption of the *attribution theory* is helpful because it asserts that people's attributions about the origin of a certain conduct contributes to influencing their emotional response to that behavior (Weiner, 1986). Causal attributions can be thus classified into

three secondary dimensions: *stability* (is the persons' behavior caused by stable or temporary factors?), *controllability* (is the persons' behavior caused by controllable or uncontrollable factors?) and *locus* (is the persons' behavior caused by internal or external factors?).

The role of teachers' attribution in modulating classroom daily responses has been well explored in educational studies: teachers' attributions predicted emotional and cognitive responses to behavior exhibited by pupils (Lovejoy, 1996; Poulou & Norwich, 2002) as well as predicting students' failure or success (Wood & Benton, 2005).

The main problem with the adoption of attributional personal theories to explain others' behaviors is linked with errors and biases that affect teachers' (and in general people's) perceptions (Heider, 1958). Such biases (i.e. *fundamental attribution error*) occurred when behavior of other people was explained by internal conditions, enduring states and personality variables rather than environmental factors that may contribute to generating the behavior itself (Heider, 1958).

Third, researchers' attention should be drawn to the multitude of factors that affect teachers' perceptions of students' challenging behaviors and consequently the quality of their relationships. Apart from the characteristics of the students such as economic background, race, age and gender (Dulin, 2001; Kokkinos, Panayiotou, & Davazoglu, 2005), other sources of "misperception" of challenging behaviors are gender stereotypes (Langfeld, 1992).

Along with students' demographics, teachers' characteristics play an active role in defining the appraisal of pupils' misconducts. For instance, less experienced teachers may express their concerns about aggressive and destructive behaviors more than their more experienced colleagues (Borg & Falzon, 1998; Kokkinos, Panayiotou, & Davazoglu, 2004). Moreover, teachers' level of occupational stress and burnout contribute to the process of labeling a given behavior as challenging. In fact, a teacher beset by a high level of fatigue or burnout may develop negative feelings toward students more than less stressed instructors (Schamer & Jackson, 1996), with the result that stressed teachers tend to react more negatively to students' misconducts.

Taking previous reflections into consideration, the teacher-student relationship was compared to a complex system shaped by several internal and external factors that contribute to determining the outcome of those interactions and that, to be studied, need the adoption of different research methodologies. In conclusion, terms such as "virtuous circles" or "vicious circles", "well behaving student" or "badly behaving student" should be thus read in light of the network of feedback and feed-forward that result from the above-mentioned "behavioral discourse".

## A TOOL FOR QUALITATIVE ANALYSIS: WORD CORRESPONDENCE ANALYSIS

Since one of our aims was to explore classroom students' challenging behaviors through the process of restructuring the relationship between words used by the teachers in describing them, textual correspondence analysis represented an appropriate and useful technique due to the presence in our study of psychometric quantitative data also.

Correspondence analysis (CA) is a multidimensional data analytic method originally developed by Benzécri (1992) during the mid-sixties aimed at describing large contingency tables and binary data that allows analysts to avoid excessive loss of information during their studies (Dore et al., 1996). Anuradha and Urs (2007) remarked that CA has several advantages over other methods of analysis: it was specifically designed to compare profiles; it was a multidimensional method that achieves appropriate data reduction, filters out noise, and objectifies correlations among variables. As a consequence of all these features, CA gained fame as a necessary tool for data analysis in nearly all disciplines (Beh, 1999).

From a strictly quantitative viewpoint, CA was guided by a pathway starting from the definition of a grouping variable and moving towards the identification of the relationship among the grouping variable and other variables under study through a mathematical model capable of describing the network of relationships underlying the phenomenon under investigation. Unfortunately, it was very unlikely that qualitative data (i.e. teachers' production of words) alone would be suitable for this kind of analysis.

At the same time, the semantic salience of qualitative data in understanding complex social phenomena was undeniable as long as researchers adopt reliable methodologies that allow them to obtain replicable results. In our case, verbal data were inspected to reveal such data structures that could give a specific idea of the key elements of the challenging behavior phenomenon as described by teachers' words. In other words, and in order to "objectively" inspect teachers' responses to open-ended questions, we chose an instrument (i.e. CA) allowing, in a synthesized and concise way, "hidden" information contained in the data to be revealed. Correspondence analysis provides very conservative data re-organization to reveal existing underlying structures using a strong graphical solution along with a strong mathematical approach to interpret results. More precisely, CA typically generates numerical indices (i.e. eigenvalues, coordinates on principal axes, squared cosine values, etc.) from which a set of low-dimensional maps can be built. Maps represent one form of visual communication that synthetically depicts the degree of association among objects under study.

In order to facilitate comprehension of how CA works, the next part of the paper is devoted to a brief explanation of the technique.

Like other multivariate statistical techniques, CA is useful to explore cross-tabulated data and social meaning variables (Harcourt, 2002). The main idea behind the technique is to build a contingency table composed of categorical variables (grouping variables) and then represent it in a more suitable two-dimensional graph on a Cartesian space. This operation allows the researcher to show the relationships between and among textual units under analysis (i.e. answers to open-ended questions).

The contingency table (in the case of qualitative data also called an occurrence table) can be described as a matrix  $M$  whose dimension is given by the number of rows ( $r_i$ ) and columns ( $c_j$ ) used to organize data. The  $(r_i, c_j)^{\text{th}}$  entry of the matrix  $M$  contains the number of times that a given row-unit (word) occurs in association with a determined column-category of the adopted variable. The next step consists of comparing each row *profile*  $(r_i)^{\text{th}}$  with the hypothetical average distribution where each set of frequency is divided by their total (in a similar fashion each column profile can be computed). The main problem is now to evaluate whether those differences are large enough to reject the hypothesis that the row frequency distribution is homogeneous: are the discrepancies between observed and expected values so large that it is unlikely they could have been generated by chance?

The comparison between the real and hypothetical distribution is calculated using the  $\chi^2$  value for each  $(r_i, c_j)^{\text{th}}$  entry of the matrix:

$$\chi^2 = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

The larger the  $\chi^2$  value, the more dissimilar the observed and expected frequencies are. In order to assess whether the  $\chi^2$  value is large enough to reject the hypothesis that the distribution of profiles is affected by chance, researchers need to compare  $\chi^2$  in association with its degrees of freedom [i.e.  $df = (r_i - 1)(c_j - 1)$ ] and the corresponding  $p$ -value. A small enough  $p$ -value indicates rejection of the null hypothesis.

The last step involves the computations of *inertia*, a measure of how much variance exist in the table; inertia is computed as follows:

$$\delta^2 = \frac{\chi^2}{N} = \frac{1}{N} \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

where  $N$  represents the total number of cross-tabulated cases (i.e. words under analysis in textual correspondence analysis).

Inertia is a crucial concept in CA in two main aspects. First, it is used to compute the Euclidean dis-

Table 1

Transformation of replies to open-ended questions in a rectangular matrix (example)

Case number	W <sub>1</sub>	W <sub>2</sub>	W <sub>3</sub>	W <sub>4</sub>	W <sub>5</sub>	W <sub>6</sub>	W <sub>7</sub>	W <sub>m</sub>
C014	hyperkinetic							... ..
C023	hyperkinetic	aggressive						... ..
C034		aggressive	provoking					... ..
C208	hyperkinetic			lack attention				... ..
C280	hyperkinetic	aggressive					swearing	... ..
C401		aggressive				ill-mannered		... ..
	...	...	...	...	...	...	...	... ..
C <sub>n</sub>	...	...	...	...	...	...	...	... ..

tance and to plot profiles in n-dimensional space, facilitating the interpretation of results. Second, global inertia can be decomposed along principal axes (which in CA correspond to the concept of *factors* in factor analysis), and it can also be decomposed amongst the row profiles or column profiles to understand which elements are “large contributors” or “small contributors” to the overall inertia. In this case, the values are labeled *row inertia* and *column inertia*, which are numbers easier to judge if compared to global inertia because they are expressed as proportions, percentages and more conveniently permills (Greenacre, 2007).

## PARTICIPANTS AND PROCEDURE

### SAMPLE

The respondents were teachers from primary and lower secondary school ( $N = 518$ ) in the city of Milan and in other urban and sub-urban areas of the Lombardy Region, Italy. Since our data were collected in the north-western part of Italy, and on the basis of the availability of school directors, our sample is just a convenience sample, and it does not claim to be representative of the whole population of Italian teachers. However, the distribution of genders in our sample closely resembles the distribution in the population of Italian teachers: in primary schools, our sample has 93.90% women and 6.10% men (in the population the percentages are 95.60 and 4.40 respectively), and in lower secondary education 78.60% were women and 21.40% men (in the population, the percentages are 76.50 and 23.50) (Italian Ministry of Education, 2006).

The number of teaching years ranged from 1 to 40 years, with an overall mean of 16.30 ( $SD = 10.50$ ).

Mean number of teaching years for men was 13.70 ( $SD = 9.40$ ) and for women 16.60 ( $SD = 10.60$ ) (see Table 1).

In Italy, there is a system of comprehensive education; there are no special education schools. Students with different abilities attend normal schools, but are helped by “special education teachers”, in a one-to-one relationship. For such reasons, data from “special education teachers” have been excluded from the present analysis.

### MATERIALS

Data were collected as part of a larger quali-quantitative international comparative study involving seven different countries in four continents.

Before expressing their “quantitative opinions”, teachers were also asked to answer some open-ended questions in order to account for their first-hand impressions. Since we are interested in common sense representation of the *challenging student*, only the first question (labeled P2AQ01 in our dataset), “Describe the behavior of the student you find most challenging in your class”, will be taken into consideration.

The quantitative facet of the research was composed of the Challenging Student Standard Questionnaire (van der Worf & Everaert, 2003), a measurement tool aimed at assessing the impact of classroom students’ challenging behaviors on teachers’ occupational stress. In responding to the questionnaire, teachers were to think of the most behaviorally challenging student in their present classroom and, keeping that student in mind, they were then asked to rate different students’ behavioral descriptors, spread over six different scales: Against the grain, Full of activity/Easily distractible, Needs a lot of attention/Weak student, Easily upset, Failure syndrome/Ex-

cessively perfectionist, and Aggressive/Hostile. Each item had to be rated twice: once for the incidence with which a behavior occurred, and once for the consequent perceived stress. The scores to the questionnaire were used to quantify the most challenging behavior the respondent had in mind during the process of answering the open-ended question. This process allowed us to explore respondents' verbal production more precisely during CA: the quantitative data provided the information (grouping variable) to assign the respondent to a given type of challenging behavior, while the qualitative data provide a deeper insight about the behavioral empirical manifestation of the same behavior analyzed through the words of those teachers coping with it.

#### PRE-PROCESSING METHODS OF LEXICAL UNITS

As already mentioned, the analysis is based on a slightly more complex tool for textual analysis (i.e. word occurrence analysis) with subsequent treatment with algorithms of factor analysis.

In order to give an idea of the technique, we have taken all answers from the questionnaires and arranged them in a data matrix in which each row-vector is one answer. In other words, to give some real examples (in which the number of the case refers to the original numbering in our database), we had data of this kind: Case 014 (C014) Hyperkinetic; Case 023 (C023) Aggressive, hyperkinetic; Case 034 (C034) Aggressive, provoking; Case 208 (C208) Hyperkinetic, he lacks attention (and so on).

We then transformed these sentences in a matrix like the following (see Table 1).

To obtain this result, we first had to *lemmatize* our words (e.g. *lacks* becomes *lack*) and, following this procedure, it was possible to build a data matrix where 1 and 0 stand for *presence* or *absence* of a certain word in the answer of a given respondent (see Table 2).

This table can in turn be transformed into a square matrix (see Table 3) in which rows and columns are single words, and each cell contains the number of co-occurrences of  $w_i$  and  $w_j$ , i.e. the number of times a certain pair of words is used in phrases of our textual corpus.

Co-occurrence tables, such as Table 3, are useful when data analysts are interested in exploring a given set of lexical production without adopting unified criteria to organize data. In such cases, the study of *row* and *column profiles* is limited to the association within words (and its corresponding representations of the Cartesian space), because the matrix  $M$  is a square one and thus *row profiles* exactly replicate *column profiles*.

In line with the objective of our research, the adoption of a variable to organize the lexical production of Italian teachers appeared a reasonable decision. Our last step was to create a contingency table in which columns represented words and rows represented a given modality of the selected grouping variable. In this case we adopted the variable "dominance scale" that is the most often occurring challenging behavior as recorded in the quantitative part of the questionnaire. In order to gain a deeper insight into the procedure, to compute the "dominance scale" in case C014 we obtained the following subscale values for quantitative scores: Against the grain = 1.30, Full of activity/Easily distractible = 3.00, Needs a lot of attention/Weak student = 2.10, Easily upset = 0.50, Failure syndrome/Excessively perfectionist = 0.90, and Aggressive/Hostile = 1.90); as a consequence, the most challenging behavior managed by the teacher will be Full of activities ( $M = 3.00$ ). The "dominance scale" values will be labeled as "2", and the description given in the open-ended question will refer to that misconduct.

Data can thus be re-organized in manner presented in Table 4.

Table 4 is exactly analogous to the contingency table we used to compute the global inertia of our data.

Table 2  
*Transformation in a rectangular co-occurrence matrix (example)*

Case number	$W_1$	$W_2$	$W_3$	$W_4$	$W_5$	$W_6$	$W_7$	...	$W_m$
C014	1	0	0	0	0	0	0	...	...
C023	1	1	0	0	0	0	0	...	...
C034	0	1	1	0	0	0	0	...	...
C208	1	0	0	1	1	0	0	...	...
C280	1	1	0	0	0	0	1	...	...
C401	0	1	0	0	0	1	0	...	...
...	...	...	...	...	...	...	...	...	...
$C_n$	...	...	...	...	...	...	...	...	...

Table 3  
Transformation in a square co-occurrence matrix (example)

	W <sub>1</sub>	W <sub>2</sub>	W <sub>3</sub>	W <sub>4</sub>	W <sub>5</sub>	W <sub>6</sub>	W <sub>7</sub>	W <sub>m</sub>
W <sub>1</sub>	–	2	0	1	1	0	1	... ..
W <sub>2</sub>	2	–	1	0	0	1	1	... ..
W <sub>3</sub>	0	1	–	0	0	0	0	... ..
W <sub>4</sub>	1	0	0	–	1	0	0	... ..
W <sub>5</sub>	1	0	0	1	–	0	0	... ..
W <sub>6</sub>	0	1	0	0	0	–	0	... ..
W <sub>7</sub>	1	1	0	0	0	0	–	... ..
	...	...	...	...	...	...	...	...
W <sub>m</sub>	...	...	...	...	...	...	...	–

Table 4  
Lemma X dominant scale variable occurrence table (example)

	W <sub>1</sub>	W <sub>2</sub>	W <sub>3</sub>	W <sub>4</sub>	W <sub>5</sub>	W <sub>6</sub>	W <sub>m</sub>
DS <sub>1</sub>	23	2	7	2	1	0	... ..
DS <sub>2</sub>	3	15	3	7	12	7	... ..
DS <sub>3</sub>	5	11	7	9	14	3	... ..
DS <sub>4</sub>	19	24	28	2	11	7	... ..
DS <sub>5</sub>	14	0	8	5	24	15	... ..
DS <sub>6</sub>	17	16	3	18	22	8	... ..

Note. DS stands for *dominance scale*. The values are coded as follows: 1 = against the grain, 2 = full of activity, 3 = weak student, 4 = easily upset, 5 = failure syndrome, 6 = aggressive/hostile.

In this case, the matrix M had a numbers of columns equal to the number of words under analysis and a number of rows equal to the modalities of the dominance scale variable; the above-mentioned concepts of *row profiles*, *row inertia* and *homogeneous distribution of frequencies* are still valid. A CA performed on that kind of contingency table allowed the researcher to compare both row profiles and column profiles as well as relevant representations of the data in n-dimensional spaces.

Alternatively, we could say that we are counting the times that certain words are used by respondents grouped according to a certain category of the quantitative variable. If it can be assumed, with Osgood (1959) that: “it seems reasonable to assume that *greater-than-chance* contingencies of items in messages would be indicative of associations in the thinking

of the source” (p. 55), then numbers in Table 4 reflect the strength of the association of meanings. In our small example, the association between W1 and DS1 (against the grain) is stronger than the association between W2 and DS1; as a matter of fact, we have two modalities (DS1 and DS2) where W4 is strongly cited, and only one modality (DS4) where we find only 2 occurrences. No association is found between DS2 and W5 (in no answers do the words W5 and dominance scale “failure syndrome” co-occur).

Contingency tables like Table 4 are easily treated with the usual algorithms for factor analysis, clustering and so on (for more details of the technique, and the implications for theories of meaning, see Lancia, 2008).

Here it will suffice to say that text-based analysis is, like most scientific methods, a process of deconstruction and reconstruction. As a result of deconstruction and subsequent reconstruction processes, something is unavoidably lost:

“[t]he deconstruction process carried out by a co-occurrence analysis leaves out three features of word/sentence meaning;

- a) the reference to the extralinguistic context (or situation), that is the *indexicality* beloved by ethnomethodologists;
- b) the *sequential order* of the words within the linguistic contexts, that is text *cohesiveness* and the anaphoric processes;
- c) the semantic effects of *speech acts*, that is all the relationships between the utterances and their *enunciation processes*” (Lancia, 2008, p. 8).

But, even if such important aspects of natural speech are lost, this approach will prove very powerful to deepen our understanding.

Since our preliminary process of lemmatization and synonymization has been a very cautious and conservative one, in order to preserve as far as possi-

ble the original integrity of texts, we had the problem of understanding the exact meaning of each lemma within the original context. Meanings can in fact be very diverse from one context to another. For example, the meaning of the word *clever* is obviously quite different in the two phrases: *the student is very clever in achieving academic results* and *the student is very clever in escaping my control*. So, contexts also have to be given due consideration.

At the end of the pre-processing phase, a total of 440 responses (i.e. *elementary contexts*) to the first open-ended instruction (“Describe the behavior of the student you find most challenging in your class”) have been analyzed. The textual corpora resulted in 6459 occurrences (i.e. *tokens*), 1163 simple lexical units and, after the adoption of an adequate threshold, the total number of words tabulated in the matrix M was 215.

## RESULTS

In Table 5 results from the application of standard algorithms for CA are reported.

The first principal dimension accounts for 25.90% of total inertia (or variance), the second one for 21.50% and, taken together, they generate a 2-dimensional space that explains 47.00% of the total inertia. The explained inertia is rather high and, as Greenacre (1993) suggests, each principal inertia accounted by a CA solution should be decomposed with respect to both columns (dominance scale) and rows (lemma) contribution to improve legibility of the results.

The first principal axis is the straight line that runs closest to the profile point (in the sense of least squares) and passes through the zero point (Greenacre, 1993); thus the major decomposition of inertia is along Dimension 1 or the X-axis and Dimension 2 or the Y-axis. The decomposition of inertia through the column profile is reported in Table 6.

Table 5

*Eigenvalues, percentages and cumulative percentage of extracted dimensions*

Factor	Inertia	Percentage	Cumulative percentage
1	0.1798	25.88	25.88
2	0.1493	21.46	47.37
3	0.1463	21.06	68.42
4	0.1182	17.01	85.43
5	0.1012	14.57	100.00

Looking at Table 6, the variable dominance scale moved along the first main axis with the following two poles: the extreme “positive” point to the right is represented by “Full of activity” ( $x = 0.2298$ ) and the extreme “negative” to the left is “Weak” ( $x = -0.6285$ ) (of course “negative” and “positive” are just artifacts of the calculus, and they could be easily inverted). Taken together, “weak” (317.50%), “aggressive/hostile” (278.00%) and “against the grain” (256.60%) contribute to nearly 80.00% of the inertia of the first axis, meaning that their contribution to the generation of the X-axis is very intense. In contrast, “failure syndrome” (5.30%) and “easily upset” (35.00%) do not account much for the inertia.

In a similar fashion, the contribution of the Y-axis to the total inertia can be computed. In this case, the Y-axis segregates “easily upset” ( $y = -1.0596$ ) and “weak” ( $y = -0.3736$ ) in the lower side, from “full of activity” ( $y = 0.3545$ ) and “aggressive/hostile” ( $y = 0.2904$ ) at the upper “positive” side. Additionally, “failure syndrome” ( $y = 0.4484$ ) lies on the positive side, but its mass is very small (4.09%) and does not contribute to the generation of the dimension. More than 80.00% of the inertia on the Y-axis comes from “easily upset” (381.20%), “full of activity” (312.50%) and “aggressive/hostile” (111.60%).

Table 6

*Coordinates and relative contributions to eigenvalues\**

Dom. scale	Mass [%]	Coordinates		Contribution to inertia*		Quality (squared cosine)		Total quality [%]
		X-Axis	Y-Axis	X-Axis	Y-Axis	X-Axis	Y-Axis	
1. Against the grain	17.76	-0.4622	0.0838	256.60	10.30	0.4189	0.0138	43.27
2. Full of activity	37.71	0.2298	0.3545	107.60	312.50	0.2135	0.5081	72.16
3. Weak Student	15.03	-0.6285	-0.3736	317.50	136.90	0.4319	0.1526	58.45
4. Easily upset	5.19	-0.3547	-1.0596	35.00	381.20	0.0445	0.3968	44.13
5. Failure Syndrome	4.10	0.1648	0.4484	5.30	47.50	0.0093	0.0692	7.85
6. Aggressive/hostile	20.22	-0.5061	0.2904	278.00	111.60	0.4131	0.1360	54.91

Note. Values are expressed in permills.



The 2-dimensional map thus generated is a fairly good representation of the structure of the data (see Figure 1).

In Figure 1 some of the results of CA of the texts describing the behavior of the challenging student are plotted on two axes representing the first two factors (the above-mentioned X- and Y-axis. In the northern part of the graph there are words such as *leader*, *disrespectful*, *overbearing*, *negative*, *rude*, *bullying* and *spiteful*, opposed to the other words in the southern part, where we find *comprehension*, *scarce*, *incapacity*, *concentration* and *sensitive*. The western side is rather unpopulated, but going to the east we find *to disturb*, *lesson*, *to interrupt*, *to provoke* and *to leave* (the northern or upper part is to be understood simply as the opposite polarity to the southern or lower part of the graph, without any other implication; the same holds for the western or left and for the eastern or right part).

We then proceeded to cluster analysis as specified above, deciding to stop the analysis at a solution with five clusters. In fact, after removing the outliers (in the end, we analyzed 396 of the 440 original texts), the distribution of objects in clusters appears a reasonably even one and, moreover, the variance between clusters was high ( $S2b = 1.049$ ), while within clusters variance was low enough ( $S2w = 0.340$ ). The comparison of the two variances is acceptable ( $S2b / (S2b + S2w) = 0.756$ ).

In Figure 2 the results of the cluster analysis, plotted against the same axes as in Figure 1, are shown.

In our experience, it is useful to read these kinds of graphs using geometric figures which can be drawn between the extreme points of the representation:

in the present case, the “triangle” drawn between Cluster two (top-left), Cluster four (bottom left) and Cluster three (extreme right). Other clusters represent an intermediate position between the extremes. The most *representative*<sup>1</sup> sentences of Cluster two are: “Quarrelsome, overbearing, aggressive towards his schoolfellows, revengeful, confused in managing his homework, not respectful of the rules of common life. I must, I can do that: this is his motto”. “Frequently conflictual with both teachers and fellow students, insecure, without firm reference points, which is why he follows negative leaders in a passive way; provoking, sometimes aggressive. Aloof from most of his fellows”. “Disrespectful of rules, very egocentric, very anxious and pressing, excessively senile, intrusive, deaf to all kinds of advice or suggestion, incoherent in his behaviors”. “Insecure, egocentric, hyperactive, incapable of controlling his emotions and moods”.

On the other extremity of the axis, Cluster four contains sentences such as: “Intolerant of every form of school rule, it is difficult for him to stay in the class or in the group, big emotional disturbances, reacts aggressively towards himself and others. Normal school work is a cause of great agitation for him, he needs a one-to-one relationship to stay serene in a different environment in order to express his interests”. “The student is aggressive towards his fellows because he finds it difficult to enter relationships and to give help to others. Minimal deficit. Deficit in organizing school material and in autonomy. Limited confidence in himself”. “Absolute lack of self-esteem and acceptance of her limits. Incapable of developing positive relationships with her fellow students. Lack of reference

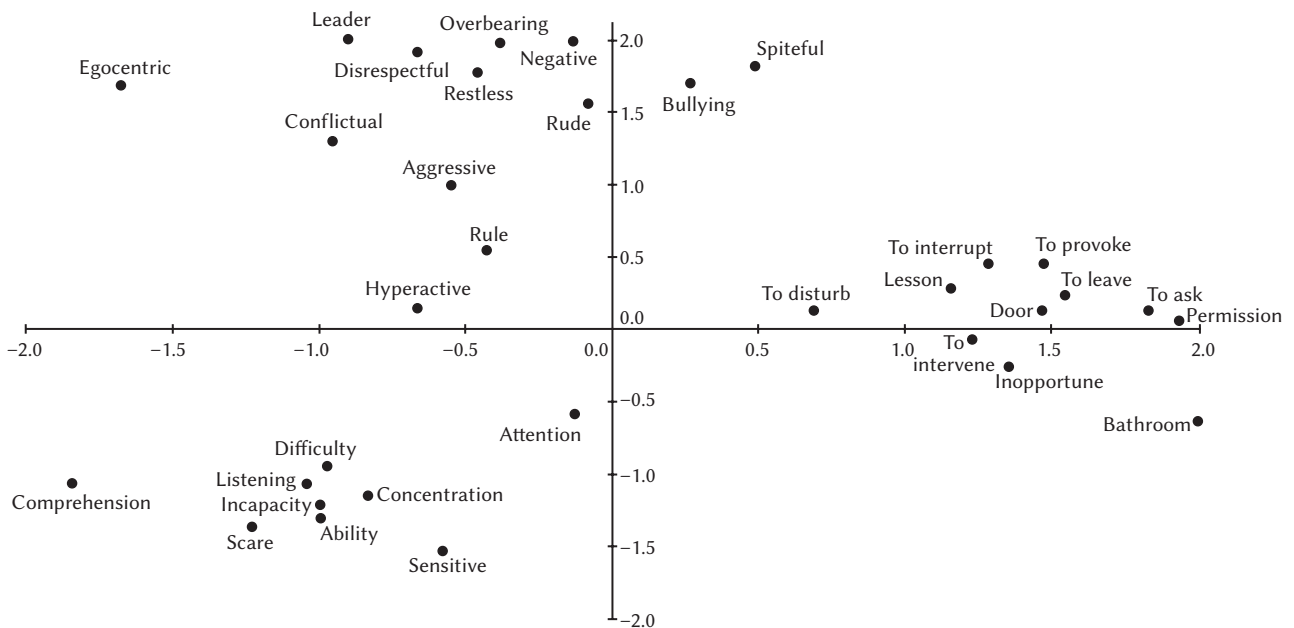


Figure 1. Graphical display of correspondence analysis of P2AQ01 “Describe the behavior of the student you find most challenging in your class”.

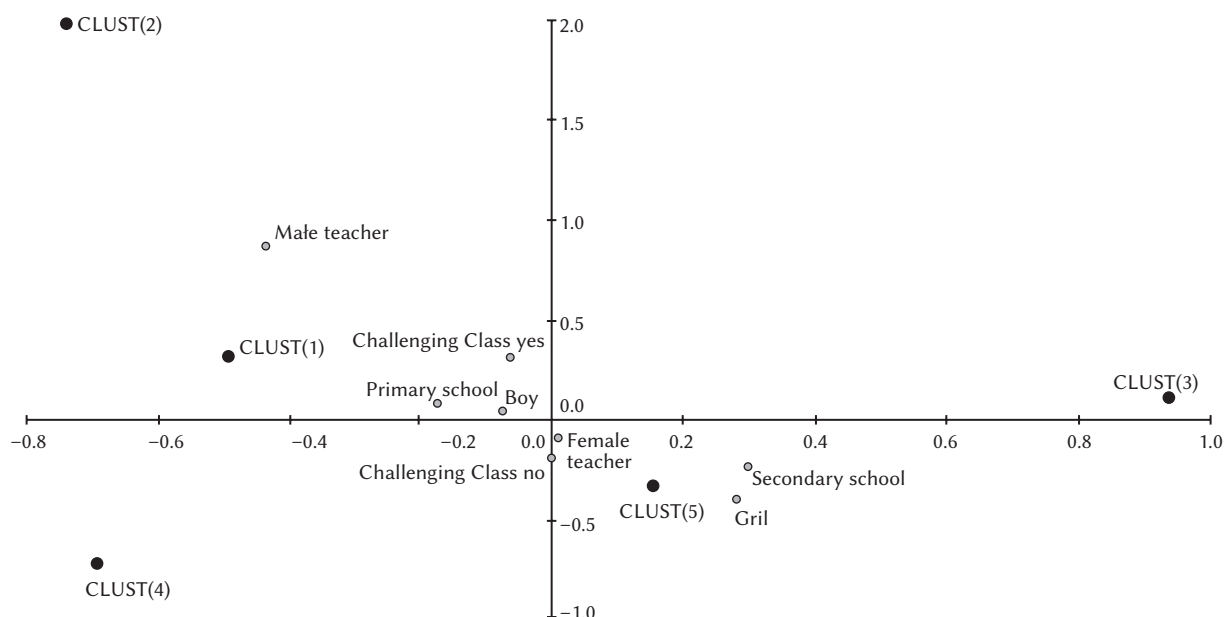


Figure 2. Graphical display of cluster analysis of P2AQ01 “Describe the behavior of the student you find most challenging in your class”.

points. Provocative and highly disturbing attitudes. Very little, if any, interest in activities”. “A foreign student, very lonely, some difficulty with the language. Deficit in relationships. He wants the attention of other fellow students. Low motivation. He is interested only in girls. He feels inferior. Lonely infancy”. “Little participation in activities, ‘absent’ during lessons, ‘closed’ in interpersonal relationships, difficulties in learning because of low application, not adequately followed by the family. The family definitely shows no interest at all in her problems”.

Interpreting the words singled out through CA in the light of the meaning of the sentences in the clusters, we can come to the conclusion that the Y-axis represents a continuum which ranges from a maximum of controllable neuroticism and antisocial attitude to a minimum represented by the very weak student, with no self-esteem, barely capable of expressing him/herself, with no interests. Under a certain point of view, we could maintain that somebody could refrain from being *disrespectful*, *overbearing*, *bullying* or *spiteful*, but cannot avoid having a *deficit*. In other words, this axis has probably something to do with the perceived *willfulness* of the student in performing a certain behavior.

In contrast, the X-axis has only one single cluster (Cluster three) at the right end. In this cluster we find sentences such as: “He likes to be the center of attention, and for this reason uses many strategies: never raises his hand to intervene, and thus expresses his opinion every time he thinks it is useful, i.e. always; disturbs lessons; he never wants to do a public oral test because he is afraid of his fellows’ evaluation; many times he does not work because he’s angry with the teacher or with other students”. “He’s

hyperactive: he jumps from the chair or does some handwork while seated (uses scissors, glue, tape...). Comments in a loud voice about completely unrelated themes. Comments about fellows. Never listens when blamed. Annoys his fellows”.

This axis obviously refers to the *disruption of teaching process*, a concept present also in the quantitative part of the questionnaire (Part 2c). And it is obvious that no objects appear in the left part of the graph: all answers report some kind of disturbance in the teaching process, and no challenging student helps the process of teaching.

## DISCUSSION

In order to explore the representation of challenging students’ behaviors through teachers’ words, the present paper reports Italian teachers’ replies to the open-ended instruction “Describe the behavior of the student you find most challenging in your class”, analyzed through a mixed-method approach.

After a long (but conservative) pre-processing phase involving the implementation of several traditional techniques in textual analysis, all available material was explored through the application of standard algorithms for word CA in order to explore the relationships amongst the most representative words used by teachers to describe the classrooms misconducts and scores in the Challenging Student Standard Questionnaire.

Results from CA of the open instruction suggested two main factors. The first factor (which we decided to label as *perceived willfulness* according to the decomposition of the inertia) presented a negative po-

larity which was very similar to the scale *Needs a lot of attention/weak student* of the quantitative part of the questionnaire. On the other end of the axis, there are other words such as *hyperactive, restless, aggressive, disrespectful, overbearing, negative, rude, bullying* and *spiteful*; it was as if our teachers cluster together three of the scales of the quantitative part, i.e. the scales of F1 *Against the grain (i)*, F2 *Full of activity/easily distractible (i)*, and F6 *Aggressive/hostile (i)*.

From a socio-interactionist approach, the first dimension (Y-axis) seems to confirm that in the appraisal of challenging behaviors, teachers' personal belief about the origin of students' misconducts play an important role. In fact, when behaviors are less perceived as "voluntary" (such as learning difficulties or being overly sensitive to moods) teachers seem to play a more care-taking role in comparison to conducts that are viewed as voluntarily destructive and aimed at deliberately undermining classroom order and discipline. In this case, the dimension segregates "easily upset" and "weak" from "full of activity" and "aggressive/hostile" students.

The second factor (X-axis) in CA can be described as the impact that the behaviors have on the *teaching process* or, using teachers' words, how the misconducts interfere with teachers' efforts to improve instructional practices and students' achievement. In this case, by the configuration of projected word profiles in CA, it can be affirmed that "aggressive/hostile" behaviors and the *weakness* (with regards to learning processes or the emotional/social aspect of schooling) of the child in the classroom exert a similar impact on weakening teachers' work and the learning process.

Our qualitative analysis seems to accurately capture even the nuances of what is portrayed by the quantitative survey (Castelli, Pepe, & Addimando, 2010; Addimando, 2010) and helps us in understanding why, in other Italian studies (Di Pietro & Ramazzo, 1997) the *Weak student* emerged as the most difficult student for Italian teachers: if this population finds it difficult to discriminate between aggressiveness, hostility and hyperactivity, putting all these behaviors in the same category, which has the *Weak student* at the other end of the scale, then it is not surprising that this last category gains salience and, maybe, over-representation in surveys. This finding can have a very important practical consequence: if we could teach teachers to discriminate between students' behaviors, and help them understand that a hyperactive student is not necessarily a hostile and aggressive one, we could have less stressed, and more effective teachers. If the teachers' training system is not equipped to provide such training, this topic could be conveniently covered by the action of a mindful and careful school psychologist.

Last not least, we believe that this work demonstrates that this kind of quali-quantitative analysis

can give a strong epistemological basis to studies on teachers' stress, being an objective analysis of inspectable datasets: our data are public, and everybody can perform the same process of deconstruction and reconstruction. We hope that these methods will contribute to give transparency and "objectivity" to psycho-pedagogical research, which is much needed these days.

#### ENDNOTE

1 "Representative" here technically means that the phrases contain the words with the highest value of the associated test measure. This measure, which is borrowed from Lebart, Morineau, and Piron (1995), has three very interesting properties: first, it correlates with absolute contributions, i.e. other measures that quantify the role played by each point of the factor space in accounting for the inertia of each factorial axis (Greenacre, 1984); second, a sign (+ or -) that can be used to order the points along a factor; third, a threshold value that we can use to reject the null hypothesis: if the value is smaller than -1.96 or larger than 1.96, we can say that the value is statistically significant ( $p \leq .050$ ). Throughout this article, answers are reported in a sequence following the ranking of their test values.

#### REFERENCES

- Addimando, L. (2010). *Comportamenti difficili degli studenti e stress degli insegnanti nelle organizzazioni educative: una ricerca internazionale* [Student's Challenging Behaviours and Teachers' Stress: an International Research]. Tesi di dottorato [Unpublished doctoral dissertation]. Università degli Studi di Milano-Bicocca.
- Anuradha, K. T., & Urs, S. R. (2007). Bibliometric indicators of Indian research collaboration patterns: A correspondence analysis. *Scientometrics*, 71, 169-179.
- Beaman, R., Wheldall, K., & Kemp, C. (2007). Recent research on troublesome classroom behaviour: A review. *Australasian Journal of Special Education*, 31, 45-60.
- Beh, E. J. (1999). Correspondence analysis of ranked data. *Communication in statistics. Theory and method*, 28, 1511-1533.
- Benzécri, J. P. (1992). *Correspondence Analysis Handbook*. New York: Marcel Dekker.
- Borg, M. G., & Falzon, J. M. (1998). Secondary school teachers' perception of students' undesirable behaviors. *British Journal of Educational Psychology*, 68, 67-79.
- Burden, P. R. (1995). *Classroom management and discipline. Methods to facilitate cooperation and instruction*. New York: Longman.

- Castelli, S., Pepe, A., & Addimando, L. (2010). Qualitatively mapping a research front through word-correspondence textual analysis: a case study. In: S. Bolasco, I. Chiari, & L. Giuliano (eds.), *Statistical Analysis of Textual Data* (pp. 677-684). Milano: LED.
- Dalenius, T. (1985). Relevant official statistics. *Journal of Official Statistics*, 1, 21-33.
- Di Pietro, M., & Ramazzo, L. (1997). *Lo stress dell'insegnante* [Teacher's Stress]. Trento: Erickson.
- Diamond, C. (1991). Knowing tornadoes and other things. *New Literary History*, 22, 1001-1015.
- Dore, J. C., Ojasoo, T., Okubo, Y., Durand, T., Dudoignon, G., & Miquel, J. F. (1996). Correspondence factor analysis of the publication patterns of 48 countries over the period 1981-1992. *Journal of American Society for Information Science*, 47, 588-602.
- Dulin, J. M. (2001). Teacher ratings of early elementary students' social-emotional behavior. *Dissertation Abstracts International*, 61, 3469.
- Emerson, E. (1995). *Challenging behaviour: analysis and intervention in people with severe intellectual disabilities*. London: Cambridge University Press.
- Grasser, A. (2009). Inaugural editorial for journal of educational psychology. *Journal of Educational Psychology*, 101, 259-261.
- Greenacre, M. J. (2007). *Correspondence Analysis in Practice* (2<sup>nd</sup> ed.). London: Chapman & Hall/CRC.
- Greenacre, M. J. (1984). *Theory and applications of correspondence analysis*. New York: Academic Press.
- Greenacre, M. J. (1993). *Theory and applications of correspondence analysis*. London: Academic Press.
- Harcourt, B. (2002). Measured Interpretation: Introducing the Method of Correspondence Analysis to Legal Studies, Symposium on Empirical and Experimental Methods in Law. *University of Illinois Law Review*, 4, 979-1017.
- Heider, F. (1958). *The Psychology of Interpersonal Relations*. New York: Wiley.
- Jones, K., Charlton, T., & Wilkin, J. (1995). Classroom behaviours which first and middle school teachers in St Helena find troublesome. *Educational Studies*, 21, 139-153.
- Kokkinos, C. M., Panayiotou, G., & Davazoglou, A. M. (2005). Correlates of teacher appraisals of student behaviors. *Psychology in the Schools*, 42, 79-89.
- Kokkinos, C. M., Panayiotou, G., & Davazoglou, A. M. (2004). Perceived seriousness of students' undesirable behaviors: The student teachers' perspective. *Educational Psychology*, 24, 109-120.
- Lancia, F. (2008). Word co-occurrence and similarity in meaning. Some methodological issues. In: S. Salvatore, & J. Valsiner (eds.), *Mind as Infinite Dimensionality*. Roma: Edizioni Carlo Amore.
- Langfeldt, H. P. (1992). Teachers' perceptions of problem behaviors: A cross-cultural study between Germany and South Korea. *British Journal of Educational Psychology*, 62, 217-224.
- Langfeldt, H. P. (1992). Teachers' perceptions of problem behaviors: A cross-cultural study between Germany and South Korea. *British Journal of Educational Psychology*, 62, 217-224.
- Lebart, L., Morineau A., & Piron M. (1995). *Statistique exploratoire multidimensionnelle* [Multidimensional Exploratory Statistics]. Paris: Dunod.
- Leung, J. P., & Ho, C. L. (2001). Disruptive classroom behavior perceived by Hong Kong primary school teachers. *Educational Research Journal*, 16, 223-237.
- Liebertson, S. (1992). Einstein, Renoir, and Greeley: Some Thoughts About Evidence in Sociology. *American Sociological Review*, 57, 1-15.
- Lorenzi-Cioldi, F. (1996). *Metodologia per la ricerca psicosociale* [Methodology for Psychosocial Research]. Milano: Il Mulino.
- Lovejoy, M. C. (1996). Social inferences regarding inattentive, overactive and aggressive child behavior and their effects on teacher reports of discipline. *Journal of Clinical Child Psychology*, 25, 33-42.
- Miller, A., Ferguson, E., & Byrne, I. (2000). Pupils' causal attributions for difficult classroom behaviour. *British Journal of Educational Psychology*, 70, 85-96.
- Osgood, C. K. (1959). The representational model and relevant research methods. In: I. de Sola Pool (ed.), *Trends in content analysis* (pp. 33-88). Urbana: University of Illinois Press.
- Poulou, M., & Norwich, B. (2002). Cognitive, emotional and behavioural responses to students with emotional and behavioural difficulties: A model of decision making. *British Educational Research Journal*, 28, 111-138.
- Schamer, L. A., & Jackson, M. J. (1996). Coping with stress. Common sense about teacher burnout. *Education Canada*, 36, 28-31.
- Schuman, H., & Presser, S. (1981). *Questions and Answers in Attitude Surveys: Experiments in Question Form, Wording, and Context*. New York: Academic Press.
- Soodak, L. C., & Podell, D. M. (1994). Teachers' thinking about difficult-to-teach students. *Journal of Educational Research*, 88, 44-51.
- Tillema, H. H. (1995). Changing the professional knowledge and beliefs of teachers: A training study. *Learning and Instruction*, 5, 291-318.
- Turnuklu, A., & Galton, M. (2001). Students' misbehaviors in Turkish and English primary classrooms. *Educational Studies*, 70, 291-305.
- Van Petegem, K., Aelterman, A., Van Keer, H., & Rosseel, Y. (2008). The influence of student characteristics and interpersonal teacher behaviour in the classroom on student's wellbeing. *Social Indicators Research*, 85, 279-291.

- Weiner, B. (1986). *An attributional theory of motivation and emotion*. New York: Springer-Verlag.
- Weiner, B. (1992). *Human motivation: Metaphors, theories, and research*. Newbury Park, CA: Sage Publications.
- Weisz, J. R. (1989). Culture and the development of child psychopathology: Lessons from Thailand. In: D. Cicchetti (ed.), *The emergence of a discipline: Rochester symposium on developmental psychopathology* (pp. 89-117). Hillsdale, NJ: Erlbaum.
- Weisz, J. R., Weiss, B., Han, S. S., Granger, D. A., & Morton, T. (1995). Effects of psychotherapy with children and adolescents revisited: A meta-analysis of treatment outcome studies. *Psychological Bulletin*, 117, 450-468.
- Wolf van Der, K., & Everaert, H. (2003). Challenging parents, teacher characteristics and teacher stress. In: R. A. Martinez-Gonzalez, M. H. Perez-Herrero, & B. Rodriguez-Ruiz (eds.), *Family school community partnerships: merging into social development*. Oviedo: Grupo SM.
- Wood, J., & Benton, S. (2005). Attributional responses to students with attention-deficit hyperactivity disorder who fail. *Teacher Education and Special Education*, 28, 153-162.
- Zimmerman, R., Khoury, E., Vega, W., Gil, A., & Warheit, G. (1995). Teacher and parent perceptions of behavior problems among a sample of African American, Hispanic, and non-Hispanic white students. *American Journal of Community Psychology*, 23, 181-197.