DOI: 10.26387/bpa.281.2

Teachers' sense of responsibility for educational outcomes. A study on the measurement properties of the teacher responsibility scale in Italian primary and secondary school teachers

Maria Cristina Matteucci, Carlo Tomasetto

Department of Psychology, University of Bologna

ABSTRACT. Lo studio presenta la traduzione italiana della Scala di Responsabilità dell'Insegnante (*Teacher Responsibility Scale*, *TRS*) in un campione di insegnanti di scuola primaria e secondaria (N = 506). Lo strumento, basato su un modello multidimensionale di responsabilità dell'insegnante, comprende quattro sottoscale che valutano la responsabilità per la motivazione degli studenti, per i risultati degli studenti, per i rapporti con gli studenti e per l'insegnamento. I risultati delle Analisi Fattoriali Confirmatorie (CFA Confirmatory Factor Analysis) supportano la struttura a quattro fattori della versione tradotta del TRS, con un'adeguata affidabilità per tutte le sottoscale e invarianza metrica del TRS per insegnanti di scuola primaria e media rispetto a insegnanti di scuola superiore. Il TRS italiano risulta quindi essere uno strumento affidabile e valido per valutare la responsabilità personale degli insegnanti per i risultati educativi.

SUMMARY. The study explored the measurement properties of an Italian translation of the Teacher Responsibility Scale (TRS) in a sample of primary and secondary school teachers (N = 506). The instrument, based on a multidimensional model of teacher's responsibility, includes four subscales assessing responsibility for student motivation, student achievement, relationships with students, and teaching. Results from a series of Confirmatory Factor Analyses (CFA) support the hypothesized four-factor structure of the back-translated version of the TRS, with adequate reliability for all subscales, and the metric invariance of the TRS for primary and middle school teachers compared to high school teachers. The Italian TRS appears to be a reliable and valid instrument to assess teachers' personal responsibility for educational outcomes, both in basic and applied research in teacher evaluation, as well as in the internal school evaluation processes.

Keywords: *Teacher responsibility, Validation, Responsibility*

INTRODUCTION

The topic of teacher responsibility is directly consequent to the personal responsibility construct, which has been studied from different perspectives and has been object of several theories from the origins of psychology (Fincham & Jaspars, 1980; Hamilton, 1978). In educational literature, the extant research focuses on two sides of the responsibility construct: teachers' collective responsibility for students' outcomes – i.e., school-level teachers' expectations and beliefs about their shared responsibility (Halvorsen, Lee & Andrade, 2009; Lee & Smith, 1996), and teachers' personal sense of responsibility, i.e., teacher's self-ascriptions of responsibility for a broad range of student needs and outcomes.

As for teacher's personal feelings of responsibility, the focus of this study, two recent contributions (Lauermann & Karabenick, 2011, 2013) concluded that the extant research was biased by conceptual ambiguity, as teacher responsibility was conceptualised as strictly intertwined with the locus of control and the self-efficacy theories. According to Lauermann and Karabenick (2013), existing empirical studies on teacher responsibility have also severe methodological limitations, as several measures of teacher responsibility have been therefore applied (e.g. measures of generic responsibility vs domain specific responsibility). As a result, the same authors have outlined the rationale for, and developed a new scale to measure how teachers view their responsibilities for educational outcomes: the *Teacher Responsibility Scale* (*TRS*). To date, this new scale has been developed and validated with American and European (i.e. German) pre-service and in-service teachers (Lauermann & Karabenick, 2013). Following the development of the TRS, teachers' personal sense of responsibility has been identified as an influencing variable in teachers' motivation and engagement and also in teaching strategies, such as instructional practices. Specifically, teacher responsibility has been identified as a key characteristic of effective teachers, with critical implications for effective instruction practices (Daniels, Radil & Wagner, 2016). Recent findings also suggest that personal responsibility predicts interest in professional development, personal time investment and work engagement (Lauermann et al., in press; Matteucci, Guglielmi & Lauermann, 2017).

The main purpose of the present work is to contribute to the Italian adaptation and validation of the TRS developed by Lauermann and Karabenick (2013). Accordingly, we will briefly outline the conceptual development of the responsibility topic in literature, mostly in relation to teachers and the educational context, subsequently, we will present an empirical investigation aiming at studying (a) reliability, (b) factorial validity, and (c) measurement invariance of the Italian version of the TRS in a sample of Italian teachers.

Theoretical perspectives on responsibility: from origins to recent educational research

From the origins of psychology, responsibility has been the focus of several theoretical perspectives which have distinguished - theoretically and empirically- causality from responsibility, and both of those concepts from blame (Shaver, 1985). Other authors (Harmon, 1995; Witt, 2001), have recognized multiple dimensions of responsibility, i.e.: agency, accountability and obligation. A multidimensional model of responsibility was developed by Lenk (Lenk & Maring, 2001) to study responsibility distribution with respect to the use of expert and information systems. According to Lenk's model, someone is responsible for something, in view of an addressee, under supervision or judgment of a judging or sanctioning instance, in relation to a criterion of attribution of accountability, within a specific realm of responsibility and action. Lauermann and Karabenick (2011) have extended this definition to education and psychological realms, by transforming it into six questions on each component of responsibility cited in this definition and discussing each of them in relation to the teaching profession and the educational context: (a) Who is responsible?, (b) For what?, (c) For/to whom?, (d) Who is the judge?, (e) In relation to what criteria of responsibility? and (f) in what realm of responsibility?, The meticulous analysis of the authors led them to phrase a definition of teacher personal responsibility as "a sense of internal obligation or commitment to produce or prevent designated outcomes or that these outcomes should have been produced or prevented" (2011, p. 15). With this original definition the authors clearly differentiated personal responsibility (a) from formal accountability, which refers to compliance with regulations, adherence to professional norms, and attainment of outcomes (Anderson & International Institute for Educational Planning, 2005), (b) from causality, which refers to the actor's causal contribution to the production of an effect or an outcome (Shaver, 1985; Lagnado & Channon, 2008), and (c) from agency, defined as

"the experience of being in control both of one's own actions and, through them, of events in the external world" (Haggard & Tsakiris, 2009, p. 52). Moreover, the Lauermann's and Karabenick's definition describes responsibility as a feeling which may both anticipate or follow a specific outcome, thus merging the sense of responsibility (Shaver, 1985) and the sense of obligation (Witt, 2001) in a unique definition.

Empirical research on teacher responsibility has mainly focused on teachers' responsibility for their students' school outcomes (Guskey, 1981, 1982; Matteucci, Tomasetto, Selleri & Carugati, 2008), however, teaching is a complex and challenging endeavour which requires professional skills and abilities and has direct implications on students (Matteucci, Carugati, Selleri, Mazzoni & Tomasetto, 2008). In this vein, a recent study (Lauermann, 2014) with elementary and high school teachers identified seven general areas of teacher responsibility, i.e.: teaching-related activities (e.g., prepare high quality lessons), student outcome (e.g., student learning and engagement), interaction with students (e.g., fairness, being a role model), positive classroom climate (e.g., create a comfortable and orderly classroom environment), interactions with others involved in students' education (e.g., parents, administration, and other teachers), school policies and external regulations (e.g., following state and district standards), as well as other general responsibilities (e.g., punctuality).

The increasing focus on teacher responsibility and the relevance of the topic is also recognizable in the latest European guidelines concerning teachers, which claims that "promoting teacher agency, empowerment and responsibility" (European Commission, 2013, p. 26) is a feasible way in the direction of strengthening the profile of the teaching profession and, therefore, supporting them to deliver higher quality teaching and to deal with complex classroom realities.

Before and besides the conceptual clarification offered by Lauermann and Karabenick (2011, 2013, 2014), educational research has linked teachers' sense of their own professional responsibility - collective and personal - to desirable and relevant outcomes. For example, teachers' collective responsibility for student learning has been associated with high student achievement gains (Lee & Smith, 1996). Therefore, responsible teachers have shown to have high expectations for all their students' learning, encouraging students and focusing mainly on positive versus negative aspects of their students (i.e. on the knowledge and skills children brought to school, rather than what they were lacking).

Teachers who ascribe themselves more responsibility for academic achievement also consider themselves as more able to influence antecedents of academic failure (Matteucci, 2007, 2008), manifest positive affect toward teaching (Guskey, 1984), and are more likely to implement new instructional practices (Guskey, 1988). Moreover, experienced responsibility for work outcomes has been found to also contribute to the explanation of teacher job satisfaction (Winter, Brenner & Petrosko, 2006) and to contribute to the feeling of work engagement (Guglielmi, Bruni, Simbula, Fraccaroli & Depolo, 2015).

To date teachers' sense of internal obligation - measured by the TRS - has proved to have critical implications for teacher motivation and psychological wellbeing (Eren, 2014; Lauermann & Karabenick, 2014; Matteucci et al., 2017; Richardson, Karabenick & Watt, 2014). For example, prospective teachers' career choice satisfaction and sense of personal responsibility were found to be positively and significantly related to each other (Eren, 2015) and teachers' sense of personal responsibility for educational outcomes contributes to the prediction of teachers' work engagement and professional commitment (Matteucci et al., 2017). Personal responsibility predicts higher interest in professional development and willingness to invest personal time in teaching-related tasks (Lauermann et al., in press).

Recently, initial evidence of implications emerged not only for teachers, but also for their students, as teacher responsibility was found to be positively related to mastery-oriented instructional practices (Kumar, Karabenick & Burgoon, 2015), and negatively to performance-oriented practices (Daniels et al., 2016).

On the whole, the above-mentioned research contributions, by showing several relevant implications of teacher responsibility, support the importance to deepen the knowledge on this topic and, therefore, suggest the need to develop and validate specific instruments to measure it.

Measures of teacher responsibility and the TRS

Before the development of the TRS, a vast array of operationalization of the teacher's responsibility concept had been employed, and five different typologies of measures were traced in literature (for a review, see Lauermann & Karabenick, 2013): (a) locus of control scales to assess teacher's assignment of responsibility for the successes or failures (e.g., *Teacher*

Locus of Control Scale; Rose & Medway, 1981), (b) single-item measures of responsibility (e.g. Matteucci & Gosling, 2004), (c) multi-item measures of responsibility for specific educational outcomes (Matteucci & Helker, 2018; Silverman, 2010), (d) generic measures of responsibility (e.g. *Teacher Job Satisfaction Questionnaire*; Lester, 1987), and (e) collective teacher responsibility measures (e.g. Lee & Smith, 1996). According to Lauermann and Karabenick (2013), the main limits of the above-mentioned scales relate to theoretical and methodological aspects. Firstly, responsibility is conceptually different from locus of control (LOC) as internal control does not automatically entail the sense of personal responsibility to implement actions, thus, scales which assess responsibility through LOC measures (a) are not appropriate. Similarly, other instruments are conceived to assess teachers' sense of responsibility for specific purposes (e.g. working with students with special needs; providing education on diversity) (c), and therefore are not appropriate to study teacher responsibility concerning their profession and everyday practices. Other instruments are conceived to assess different constructs (e.g. job satisfaction) and include only a few items to measure teacher responsibility as a sub-dimension (d). Similarly, from the methodological point of view, as responsibility is acknowledged to be a multi-dimensional construct, single-item measures (b) are not considered appropriate to assess it.

The review of these measures led Lauermann and Karabenick (2013) to conclude the need of a specific instrument to assess to what extent teachers feel responsible for specific aspects related to their profession. The preliminary scale items were developed through a conceptually- and empirically-driven procedure which led the authors to focus on outcome-based key responsibilities with which most teachers could identify (i.e. students' achievement, students' motivation, having positive relationships with students and teaching quality). According to a general process that, over the past century, gradually assigned the responsibility for students' academic success from students and their families to educators (Coleman, 1968), the scale assesses teachers' sense of personal responsibility for providing educational services (e.g., preparing engaging lessons in order to increase student interest), as well as for outcomes (e.g., students' low achievement, lack of interest, etc.). Accordingly, items were formulated to assess teachers' willingness to assume personal responsibility for several negative educational outcomes that they should have prevented. The design and validation procedure led to a multidimensional scale with four subscales to assess teachers' willingness to assume personal responsibility for negative hypothetical educational outcomes that can occur in any classroom at any time (see "Translation procedure" section for details about items and sub-scales and Appendix for the original scale).

The validation study of the TRS (Lauermann & Karabenick, 2013) provided evidence that the scale is applicable to both pre-service and in-service teachers, and its validity has been confirmed across the US and the German educational systems. To date, validation studies assessing the metric properties of the TRS in diverse cultural and educational settings are still limited. Translated versions of the TRS have been employed but not validated with a sample of French- speaking Swiss vocational teachers (Berger, Girardet & Aprea, 2013), German university teachers (Wosnitza, Helker & Lohbeck, 2014), and a Turkish version, which obtained good internal validity and reliability indexes with a sample of prospective teachers (Eren, 2014).

THE PRESENT STUDY

The main purpose of the present work is to contribute to the Italian adaptation and validation of the TRS developed by Lauermann and Karabenick (2013) by investigating (a) reliability, (b) factorial validity, and (c) measurement invariance of the Italian version of the TRS in a sample of Italian teachers.

METHODS

Translation procedure

The original English version of the TRS includes 12 items designed to represent the following four areas of responsibility: responsibility for student motivation; (e.g., "I would feel personally responsible if a student of mine was not interested in the subject I teach"); student achievement (e.g., "I would feel personally responsible if a student of mine had very low achievement"); relationships with students (e.g., "I would feel personally responsible if a student of mine did not think that he/she can trust me with his/her problems in or outside of school"); and teaching. (e.g., "I would feel personally responsible if a lesson I taught failed to reflect my highest ability as a teacher"). The items are preceded by the

following statement: "Imagine that the following situations would occur when you have classes of your own. To what extent would you feel personally responsible that you should have prevented each of the following?". The items are rated on a seven-point Likert scale, ranging from 0 (not at all) to 6 (completely). The original TRS was translated into Italian by means of a forward-backward-forward approach, in order to ensure the linguistic equivalence between the Italian and the original English version of the instrument.

Two independent researchers translated the original scale in Italian. The divergences emerged when the first independent steps of work were discussed, and the reconciled translation was backwardly translated in to English by a mother tongue-language expert, in order to detect possible mismatches. As no relevant problems emerged at the end of the whole procedure, the resulting Italian version of the TRS (presented in Appendix) and the English version of the TRS contain the same item and scale formatting.

Participants and procedure

The data were collected from 219 public primary (61%) and middle (39%) school teachers aged between 25 and 70 years (M = 48.86; SD = 8.36; 90.4% women) and 287 high school teachers aged between 27 and 64 years (M = 49.92; SD = 7.10; 63.5% women). The average level of teaching experience was 17.25 (SD = 11.20) for secondary school teachers and 16.09 (SD = 10.49) for primary and middle school teachers. The schools were selected across three regions in North, Centre and South of Italy. High schools were: 29.4% lyceum; 66.7% technical institutes and 3.9% professional institutes. A researcher visited each school in the three regions and presented the survey to the school principal or assistant principal. A request to fill an online questionnaire was sent to teachers of the schools who had agreed to participate via the school email system. Teachers' participation was voluntary and informed consent was obtained from each participant.

Data analyses

Internal consistency was assessed by means of Cronbach's α coefficients, zero-order inter-item correlation, and item-total correlation coefficients.

To assess the fit between the hypothesized four-dimensional structure and the observed data for the Italian version of the TRS, we conducted a Confirmatory Factor Analysis (CFA) based on a Maximum Likelihood estimation procedure, using AMOS 22 (Arbuckle, 2013). To assess the overall adequacy of the model we examined the ratio of chi-square to its degrees of freedom (χ^2/df), the Comparative Fit Index (CFI; Bentler, 1990), the Tucker-Lewis Index (TLI; Tucker & Lewis, 1973), and the Root Mean Square Error of Approximation (RMSEA; Steiger & Lind, 1980). Values of $\chi^2/df \leq 2$, CFI and TLI > .95, and RMSEA < .06 were assumed as representing good fit; values of $\chi^2/df \leq 3$, CFI and TLI > .90, and RMSEA < .08 were deemed as reflecting acceptable fit (see Hooper, Coughlan & Mullen, 2008; Schermelleh-Engel, Moosbrugger & Muller, 2003; Schreiber, Nora, Stage, Barlow & King, 2006). We also ran a supplementary CFA with a unique latent factor representing the overall construct of teachers' responsibility to assess whether a simpler factorial structure may provide a more parsimonious representation of the Italian version of the TRS.

Measurement invariance across school level-groups was examined by testing and comparing four nested models (Model 1 to Model 4) using Multi-group Confirmatory Factor Analyses (Byrne, 2004; Vandenberg & Lance, 2000). As a prerequisite for assessing the measurement equivalence of the Italian version of the TRS across school levels, the fit of the hypothesized model was established for primary and middle school teachers and high school teachers separately. Each successive model included the previous model restrictions plus additional constraints (Meredith & Teresi, 2006). We decided to compare primary and middle school teachers, on the one hand, with high school teachers, on the other, as in the Italian Education system after successful completion of primary school (level 1, according to the 2011 International Standard Classification of Education [ISCED], cfr. Schneider, 2013), all students progress to a common-track middle school (the lower secondary level, or ISCED level 2). Differently from these two levels of education, where students follow the same general common core curriculum, students alternatively enrol in secondary general or vocational education (ISCED level 3), in preparation for tertiary education or to acquire skills relevant to employment, or both (European Commission, 2014). Thus, the transition from middle school to a secondary education program marks the beginning of a completely different learning experience, and it is at that point that Italian students choose distinct educational or vocational pathways.

RESULTS

The proportion of missing item responses for each scale ranged from .9% to 4.7% and the Little's MCAR test confirmed that data were missing at random. Therefore, missing data were imputed by means of a maximum likelihood approach with the Expectation-Maximization (EM) algorithm (single imputation) (Allison, 2002; McLachlan & Krishnan, 2007).

Item means, standard deviations, and zero-order correlations among observed variables are presented in Table 1. Bivariate correlations show that, within each subscale, items are highly and positively correlated (RSM, *range*: .66-.74; RSA, *range*: .66-.74; RRS, *range*: .80-.84; RTE, *range*: .73-.90), as well as each item with the total score of the scale (*range*: .68-.84). Moderate negative correlations indexes emerged concerning the teachers' age, suggesting that teachers' sense of personal responsibility decreases gradually with age. Scale scores distributions and reliability are reported in Table 2. Cronbach's α for each of the four subscales range from .87 to .92, indicating good internal consistency.

Factorial structure of the Italian version of the TRS

To assess the fit between the hypothesized four-factor structure and the observed data for the Italian version of the TRS we conducted a Confirmatory Factor Analysis (CFA) based on Maximum Likelihood estimation procedure.

Results of the CFA revealed that the model had a good fit according to the CFI and TLI indices (.98 and .97, respectively), an acceptable fit according to the RMSEA (.06), and an inadequate fit according to the χ^2/df index (3.19). However, the inspection of modification indices suggested that the adequacy of the model could be further improved by relaxing the assumption of independence of measurement errors between some of the items. In particular, the most substantial improvement could be obtained by admitting a co-variation between the measurement errors of the RRS_1 and RRS_2 items (i.e., "*A student of mine thought he/she could not count on me when he/she needed help with something*", and "*A student of mine did not think that he/she could trust me with his/her problems in or outside of school*"). As both the items were assumed to load on the same latent factor (i.e., Responsibility for Relationships with students), and the content of the two items was strikingly similar, we decided to repeat the analysis by admitting the residuals of the two items to co-vary. Results of the CFA reveal that the modified model had a good fit to the data (χ^2/df = 2.78, CFI = .98, TLI = .97, RMSEA = .05), and should therefore be retained as a valid representation of the factorial structure of the scale (see Figure 1).

As the correlations among the four latent variables were fairly high (*range*: .62-.82), we also tested the fit of a more parsimonious model in which all the observed items were forced to load onto a unique factor (i.e., overall teacher's responsibility). However, the fit of the unique-factor alternative model was poor (χ^2/df = 23.37, CFI = .74, TLI = .68, RMSEA = .22), and did not support the existence of a single latent dimension accounting for different facets of teachers' responsibility, as measured by the TRS.

Measurement invariance across school grades

As a prerequisite for assessing the measurement equivalence of the Italian version of the TRS across school levels, the fit of the hypothesized four-factor model was established for primary and middle school teachers, and for high school teachers separately. Results confirmed that fit indexed were adequate for both groups (χ^2/df = 2.51, CFI = .97, TLI = .96, RMSEA = .07, and χ^2/df = 2.50, CFI = .97, TLI = .96, RMSEA = .07 for primary and middle school and for high school teachers, respectively).

Then, to establish the equivalence of different measurement properties across the two groups, we conducted a series of Multigroup Confirmatory Factor Analysis in which we compared a series of nested models with increasing constraints (Meredith & Teresi, 2006). In Model 1 we evaluated the fit of the baseline model, in which only the factor structure (the number of factors and the paths admitted from observed indicators to latent variables) was fixed to be equal across the two groups (*configural invariance*). In Model 2 we assessed the equivalence of the relations between each item and the corresponding latent construct by constraining all factor loadings to be equal across groups (*metric invariance*). In Model 3 we assessed the equivalence of the items' means by forcing all observed indicators' means to be equal across groups (*scalar invariance*). Finally, in Model 4 we tested whether measurement error residuals were the same for each item across groups (*measurement error invariance*). Overall fit indices were examined separately for each model. The

Table 1 – Descriptive statistics, internal reliabilities, and zero-order correlations for all measured variables (N = 506)

	M	SD	Item-subscale correlation	α	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.
1. RSM 1	4.71	1.26	.74	.87	.74**	.66**	.62**	.59**	.60**	.40**	46**	.49	.52**	.54**	.44**	.74**	–.11*
2. RSM 2	4.33	1.32	.76		-	.67**	.59**	.51**	.55**	.34**	.41**	.43**	.42**	.43**	.36**	.68	–.11*
3. RSM 3	4.38	1.35	.70			-	.56**	.54**	.57**	.48**	.52**	.53**	.52**	.48**	.44**	.73**	–.13**
4. RSA 1	4.51	1.26	.68	.87			-	.66**	.66**	.44**	.47**	.52**	.50**	.51**	.45**	.75**	–.12**
5. RSA 2	4.25	1.18	.75					-	.74**	.45**	.42**	.48**	.45**	.46**	.39**	.70**	–.12**
6. RSA 3	4.18	1.22	.74						-	.47**	.45**	.50**	.49**	.47**	.39**	.71**	–.09†
7. RRS 1	5.06	1.54	.85	.93						-	.83**	.80**	.64**	.62**	.55**	.78**	–.14**
8. RRS 2	4.99	1.53	.88								-	.84**	.65**	.65**	.59**	.81**	–.13**
9. RRS 3	5.02	1.52	.85									-	.73**	.73**	.64**	.84**	–.12**
10. RTE 1	5.25	1.36	.85	.92									-	.90**	.73**	.81**	–.15**
11. RTE 2	5.25	1.29	.88											-	.76**	.81**	–.14**
12. RTE 3	5.15	1.31	.75												-	.71**	–.11*
13. TRS tot	4.82	1.03	-	-												-	–.15**
14. Age	49.48	7.67	-	-													-

Note. $p<.10$; *$p<.05$; **$p<.01$ (two tailed).

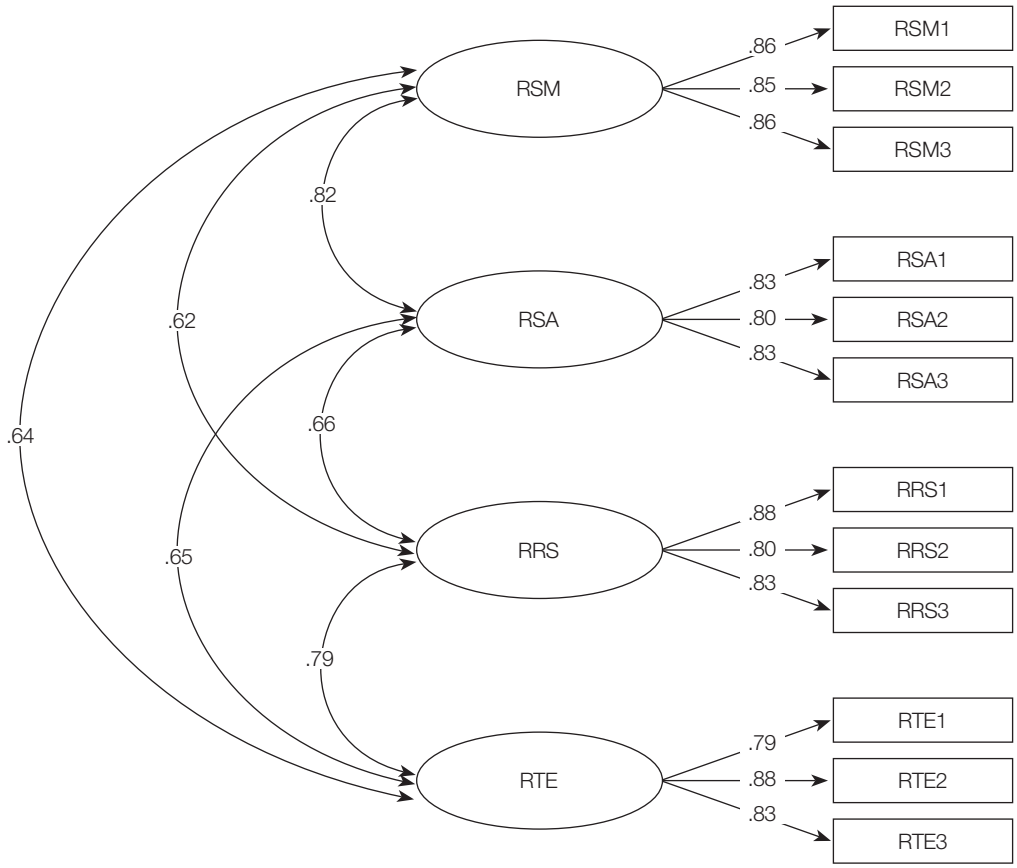
Legenda. RSM = Responsibility for Student Motivation; RSA = Responsibility for Student Achievement; RRS = Responsibility for Relationships with Students; RTE = Responsibility for Teaching; TRS response scale: *range* 1-7

Table 2 – Descriptive statistics, symmetry, kurtosis and Cronbach's alpha for each size of the TRS scale

TRS dimensions	M (SD)	Min-Max	Symmetry	Kurtosis	α
RSM	4.53 (1.11)	1-7	−.31	.17	.87
RSA	4.36 (1.03)	1-7	−.07	.10	.87
RRS	5.12 (1.35)	1-7	−.81	.22	.93
RTE	5.32 (1.09)	2-7	−.62	.09	.92

Legenda. RSM = Responsibility for Student Motivation; RSA = Responsibility for Student Achievement; RRS = Responsibility for Relationships with Students; RTE = Responsibility for Teaching; TRS response scale: range 1-7.

Figure 1 – Factorial structure of the Italian version of the Teacher Responsibility Scale



Note. Standardized coefficients are reported. All coefficients are significant at the $p<.001$ level. Residual terms are not shown in the figure. Correlation between residual terms is admitted between items RRS1 and RRS2 ($r = .32$).

criteria for assessing the difference between the competing models was the chi-square test difference, and the difference in the CFIs between competing models. A non-significant chi-square test difference, and a difference of CFI values smaller than .01(Cheung & Rensvold, 2002), were deemed as supporting the more constrained of the competing models. The overall and comparative fit statistics of invariance models are reported in Table 3.

Table 3 – Fit indices and comparison between invariant models across school grades (primary/middle schools vs high school)

	$\chi^{2\,(df)}$	χ^2/df	CFI	TLI	RMSEA	Model comparison	$\Delta\chi^2$	Δdf	*p*	CFI
Model 1	194.40 (94)	2.06	.97	.97	.04	-	-	-	-	-
Model 2	200.78 (102)	1.96	.98	.97	.04	Model 1	6.38	8	.60	.01
Model 3	271.11 (114)	2.37	.96	.96	.05	Model 2	70.32	12	< .001	–.01
Model 4	456.03 (137)	3.32	.93	.93	.06	Model 3	184.92	23	< .001	–.03

Legenda. χ^2 = chi square; df = degrees of freedom; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation; Δdf = difference in degrees of freedom between nested models; $\Delta\chi^2$ = difference between χ^2 of nested models; *p* = probability value of $\Delta\chi^2$ test; ΔCFI = difference between CFIs of nested models; Model 1 = equality of factor structure (baseline); Model 2 = Model 1 + equality of factor loadings; Model 3 = Model 2 + equality of items means; Model 4 = Model 3 + equality of error variances.

Results showed an adequate fit of Model 1, thus indicating that observed indicators reflect the same underlying constructs across the two school levels. As the chi-square test difference between Model 2 and Model 1 was not significant, and the CFI difference was lower than 1, equivalence in factor loadings was also established. Conversely, although the CFI difference was small (<.01), the chi-square test difference between Model 2 and Model 1 was significant, thus indicating that item intercepts were not equivalent across groups, and scalar invariance of the TRS for middle school and college teachers was therefore not supported. As seen in Table 4, univariate tests confirm that mean scores of middle-school teachers are systematically higher than those of college teachers on all the items. In sum, results indicated that configural and metric invariance across school levels was supported, whereas scalar and measurement error invariance were not.

DISCUSSION

The main purpose of the present work was to present the translated version of the TRS and to study (a) reliability, (b) factorial validity, and (c) measurement invariance of the Italian version of the TR in a sample of Italian teachers. The results indicate an appropriate internal consistency of the Italian version of the TRS since Cronbach's α coefficients are excellent or good and comparable with those obtained in the validation study (Lauermann & Karabenick, 2013).

Regarding factorial validity, the results support the hypothesized four-factor structure of the Italian version of the TRS, confirming the presence of four interrelated but separate dimensions, that assess responsibility for student motivation, student achievement, relationships with students and responsibility for teaching. The findings are consistent with the hypothesized structure resulting from the original validation study of the TRS with American and German teachers (Lauermann & Karabenick, 2013).

Results of the measurement invariance analysis showed an adequate fit of the baseline model (Model 1), in which only the factor structure (the number of factors and the paths admitted from observed indicators to latent variables) was fixed to be equal across the two groups (configural invariance), thus indicating that observed indicators reflected the same underlying constructs across the two school levels. Results also supported the metric invariance of the tool, confirming that the relations between each item and the corresponding latent construct was equal for teachers in different school levels. However, the scalar invariance of the TRS for primary and middle school teachers and for high-school teachers was not supported, as mean scores of primary and middle school teachers were systematically higher than those of high-school teachers on all the items.

These findings add new insight and expand prior knowledge concerning teachers' acceptance of personal

Table 4 – Mean differences between primary/middle school vs. high school teachers on TRS items (N = 495)

	Primary/middle school		High school			
	M	SD	M	SD	*t* value	Cohen's *d*
1. RSM 1	5.11	.98	4.49	1.29	5.84**	.52
2. RSM 2	4.68	1.21	4.13	1.32	4.87**	.43
3. RSM 3	4.74	1.08	4.22	1.39	4.53**	.40
4. RSA 1	4.87	1.01	4.31	1.29	5.28**	.47
5. RSA 2	4.55	1.01	4.10	1.87	4.49**	.40
6. RSA 3	4.63	1.04	3.91	1.17	7.12**	.64
7. RRS 1	5.44	1.19	4.90	1.59	4.19**	.37
8. RRS 2	5.33	1.21	4.88	1.59	3.44**	.30
9. RRS 3	5.41	1.14	4.90	1.57	4.01**	.36
10. RTE 1	5.55	1.04	5.19	1.34	3.24**	.29
11. RTE 2	5.55	1.01	5.17	1.23	3.69**	.33
12. RTE 3	5.35	1.04	5.19	1.23	1.58^{ns}	.14

Note. ns = non significant; ***p<.01 (two tailed)*.
Legenda. RSM = Responsibility for Student Motivation; RSA = Responsibility for Student Achievement; RRS = Responsibility for Relationships with Students; RTE = Responsibility for Teaching; TRS response scale: *range* 1-7.

responsibility for educational outcomes. In past works (Lauermann & Karabenick, 2013) the scale was tested with a sample of kindergarten through 12^{th} grade (K-12) regular in-service teachers, and with secondary-level pre-service teachers. To date, this is the first time the TRS has been tested with secondary-level regular in-service teachers. The lower level of responsibility that secondary-level (high school) teachers accept to hold, might, therefore be determined by the students' age: as it is compulsory to attend school until age 16, teachers of the secondary schools involved in this study work have a curriculum aimed at graduating students at age 18 (high school diploma) preparing them for a job or for master degree programs or other post-secondary education programs. As a consequence, teachers in secondary school, when answering the scale, may refer to students who take charge of their academic success, who are held accountable for achieving learning goals. In secondary education learners are expected to fulfil their responsibilities as students, and the school requires students to become responsible and accountable for their own academic success, thus shifting responsibility - at least in part - from teachers to students.

Moreover, it is to note that configurable and metric invariance are strictly necessary for basic research purposes, as they establish the fact that manifest indicators assess the same underlying construct across groups (Meredith & Teresi, 2006). Conversely, strong and strict invariance are of less substantive importance, as group differences in observed scores may reflect meaningful underlying group differences, whereas differences in residual variances may reflect differences in measurement reliability rather than in the scale validity (Vandenberg & Lance, 2000).

In conclusion, the Italian version of the TRS appears to be a reliable and valid instrument to assess teachers' personal responsibility for educational outcomes, applicable across different educational contexts, both for basic and applied research in educational psychology, as well as for intervention programs with teacher.

Limitations

In addition to our main findings, it is important to acknowledge a set of limitations and directions for future research. From the statistical point of view, the main limit concerns the scalar invariance of the TRS for primary and middle school teachers and for high-school teachers, which was not supported. However, as explained in the results discussion, strong and strict invariance may be less important in the context of research in which group differences in specific factors are indicative of individual differences.

Similarly to the original scale, probably the main limits of the scale are firstly directly linked to survey methodology based on self-report measures, which imply a certain risk that the findings may be biased by the influence of social desirability. Moreover, questionnaire-based measures - like the TRS- present a set of standard items and respondents' answers are limited to a fixed set of responses, which may prejudice the possibility to capture further dimensions of responsibility or differences in the amount of responsibility perceived by the teachers.

Secondly, as the TRS's authors acknowledged:

> "it is critical to recognize that teachers' professional responsibility is embedded in a variety of contexts; teachers may feel different degrees of responsibility depending on the characteristics of their teacher education program, their students' characteristics, school characteristics, and characteristics of the education system" (Lauermann & Karabenick, 2013, p. 24).

Further research should therefore study differences in teacher responsibility at school level (organizational culture), as well as the role of personal and contextual influences, such as, for example the role of school principal and of school collective responsibility as a whole. Longitudinal modifications on teachers' personal sense of responsibility during their careers, and cross-sectional changes in perceived responsibility according to student or classroom characteristics may also be further investigated. Finally, the use of a mixed-methods methodology, integrating quantitative and qualitative data, could help overcome the weakness of a questionnaire-based survey.

CONCLUSIONS

To date the research has provided initial evidence that personal responsibility may be a pivotal variable in order to strengthen the profile of the teaching profession. For example, the European Commission claims that promoting teacher agency, empowerment and responsibility might be an effective way to develop teacher competence (European Commission, 2013). Moreover, in order to establish positive and effective school-family relationships, a shared viewpoint on reciprocal and mutual responsibilities among teachers, as well as with parents and students, needs to be established. Therefore, the existence of a scale aiming at examining and monitoring the teachers' sense of personal responsibility is a pre-requisite to realize interventions and professional development activities aimed at promoting teachers' personal responsibility and, finally, to ameliorate learning/teaching conditions.

Disclosure statement. The authors declare that they have no conflict of interest.

References

ALLISON, P.D. (2002). Missing data: Quantitative applications in the social sciences. *British Journal of Mathematical and Statistical Psychology*, *55*, 193-196. doi: 10.1348/000711002159653

ANDERSON, J.A. & INTERNATIONAL INSTITUTE FOR EDUCATIONAL PLANNING (2005). *Accountability in education*. Paris: International Institute for Educational Planning.

ARBUCKLE, J. (2013). AMOS 22. *User's Guide*. Chicago, IL: SmallWaters Corporation.

BENTLER, P.M. (1990). Comparative fit indexes in structural models. *Psychological bulletin*, *107*, 238-246. doi: http://dx.doi.org/10.1037/0033-2909.107.2.238

BERGER, J.-L., GIRARDET, C. & APREA, C. (2013, August). *Motivation, sense of responsibility, and classroom management self-reported practices in vocational teachers*. Paper presented at the 15th Biennial Conference of the European Association for Research on Learning and Instruction (EARLI), Munich, Germany, August 27-31.

BYRNE, B.M. (2004). Testing for multigroup invariance using AMOS

graphics: A road less traveled. *Structural Equation Modeling*, *11*, 272-300. doi: 10.1207/s15328007sem1102_8

CHEUNG, G.W. & RENSVOLD, R.B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233-255. doi: 10.1207/S15328007SEM0902_5

COLEMAN, J.S. (1968). The concept of equality of educational opportunity. *Harvard Educational Review*, *38*, 7-22.

DANIELS, L.M., RADIL, A. & WAGNER, A.K. (2016). Concordance between Preservice Teachers' Personal Responsibilities and Intended Instructional Practices. *The Journal of Experimental Education*, *84*, 529-553. doi: 10.1080/00220973.2015.1054333

EREN, A. (2014). Uncovering the links between prospective teachers' personal responsibility, academic optimism, hope, and emotions about teaching: A mediation analysis. *Social Psychology of Education*, *17*, 73-104. doi: 10.1007/s11218-013-9243-5

EREN, A. (2015). 'Not only satisfied and responsible, but also hopeful': Prospective teachers' career choice satisfaction, hope, and personal responsibility. *Cambridge Journal of Education*, *45* (2), 149-166.

EUROPEAN COMMISSION (2013). *Supporting Teacher Competence Development for Better Learning Outcomes*. Retrived from http://ec.europa.eu/education/policy/school/doc/teachercomp_en.pdf

EUROPEAN COMMISSION (2014). *The Structure of the European Education Systems 2014/15: Schematic Diagrams. Resource document European Commission*. Retrived from http://eacea.ec.europa.eu/education/eurydice/facts_and_figures_en.php#diagrams

FINCHAM, F.D. & JASPARS, J.M. (1980). Attribution of responsibility: From man as the scientist to man as lawyer. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology*. New York NY: Academic Press.

GUGLIELMI, D., BRUNI, I., SIMBULA, S., FRACCAROLI, F. & DEPOLO, M. (2016). What drives teacher engagement: A study of different age cohorts. *European Journal of Psychology of Education*, *31*, 323-340. doi: 10.1007/s10212-015-0263-8

GUSKEY, T.R. (1981). Measurement of the responsibility teachers assume for academic successes and failures in the classroom. *Journal of Teacher Education*, *32* (3), 44-51. doi: 10.1177/002248718103200310

GUSKEY, T.R. (1982). Differences in teachers' perceptions of personal control of positive versus negative student learning outcomes. *Contemporary Educational Psychology*, *7* (1), 70-80. doi: 10.1016/0361-476X(82)90009-1

GUSKEY, T.R. (1984). The influence of change in instructional effectiveness upon the affective characteristics of teachers. *American Educational Research Journal*, *21* (2), 245-259.

GUSKEY, T.R. (1988). Teacher efficacy, self-concept, and attitudes toward the implementation of instructional innovation. *Teaching and Teacher education*, *4* (1), 63-69.

HAGGARD, P. & TSAKIRIS, M. (2009). The experience of agency feelings, judgments, and responsibility. *Current Directions in Psychological Science*, *18*, 242-246. doi:10.1111/j.1467-8721.2009.01644.x

HALVORSEN, A.L., LEE, V.E. & ANDRADE, F.H. (2009). A mixed-method study of teachers' attitudes about teaching in urban and low-income schools. *Urban Education*, *44*, 181-224. doi: 10.1177/0042085908318696

HAMILTON, V.L. (1978). Who is Responsible? Toward a Social Psychology of Responsibility Attribution. *Social Psychology*, *41*, 316-328. Retrieved from http://www.jstor.org/stable/3033584

HARMON, M.M. (1995). *Responsibility as paradox: A critique of rational discourse on government*. Thousand Oaks, CA: Sage

HELKER, K. & WOSNITZA, M. (2013, August). *Early Career University Teachers' Sense of Responsibility for Teaching*. Paper presented at the 15th Biennial Conference of the European Association for Research on Learning and Instruction (EARLI), Munich, Germany.

HOOPER, D., COUGHLAN, J. & MULLEN, M. (2008). Evaluating model fit: A synthesis of the structural equation modelling literature. In A. Brown (Ed.), *7th European Conference on research methodology for business and management studies*. UK: Academic Published Limited Reading.

KUMAR, R., KARABENICK, S.A. & BURGOON, J.N. (2015). Teachers' implicit attitudes, explicit beliefs, and the mediating role of respect and cultural responsibility on mastery and performance-focused instructional practices. *Journal of Educational Psychology*, *107*, 533-545. doi: http://dx.doi.org/10.1037/a0037471

LAGNADO, D.A. & CHANNON, S. (2008). Judgments of cause and blame: The effects of intentionality and foreseeability. *Cognition*, *108*, 754-770. doi:10.1016/j.cognition.2008.06.009

LAUERMANN, F. (2014). Teacher responsibility from the teacher's perspective. *International Journal of Educational Research*, *65*, 75-89. doi: http://dx.doi.org/10.1016/j.ijer.2013.09.005

LAUERMANN, F. & KARABENICK, S.A. (2011). Taking teacher responsibility into account(ability): Explicating its multiple components and theoretical status. *Educational Psychologist*, *46*, 122-140. doi:10.1080/00461520.2011.558818

LAUERMANN, F. & KARABENICK, S.A. (2013). The meaning and measure of teachers' sense of responsibility for educational

outcomes. *Teaching and Teacher Education*, *30*, 13-26. doi:10.1016/j.tate.2012.10.001

LAUERMANN, F. & KARABENICK, S.A. (2014). Teacher responsibility: What does it mean for teachers' motivation and emotions? In P.W. Richardson, S.A. Karabenick & H.M.G. Watt (Eds.), *Teacher motivation: Implications for theory and practice*. New York, NY: Routledge/Taylor & Francis.

LAUERMANN, F., KARABENICK, S.A., CARPENTER, R. & KUUSINEN, C. (in press). Teacher motivation and professional commitment: The role of motives for teaching, teacher efficacy and sense of professional responsibility. In H.M.G. Watt, P.W. Richardson & K. Smith (Eds.), *Why teach? Antecedents and Consequences across Cultural Settings*. Cambridge: Cambridge University Press.

LEE, V.E. & SMITH, J.B. (1996). Collective responsibility for learning and its effects on gains in achievement for early secondary school students. *American Journal of Education*, *104*, 103-147. Retrived from http://www.jstor.org/stable/1085702

LENK, H. & MARING, M. (2001). Responsibility and technology. In A.E. Auhagen & H.W. Bierhoff (Eds.), *Responsibility: The many faces of a social phenomenon*. London, New York NY: Routledge.

LESTER, P.E. (1987). Development and factor analysis of the teacher job satisfaction questionnaire (TJSQ). *Educational and Psychological Measurement*, *47*, 223-233. doi: 10.1177/0013164487471031

MATTEUCCI, M.C. (2007). Teachers facing school failure: Social representations of intelligence and behavioural feedback. *Social Psychology of Education*, *10*, 29-53. doi: 10.1007/s11218-006-9011-x

MATTEUCCI, M.C. (2008). Transferring corporate social responsibility to the school context: Teachers' perceived responsibility for students' failure. In A.R. Timpere (Ed.), *Corporate Social Responsibility*. New York, NY: Nova Science Publishers.

MATTEUCCI, M.C. & GOSLING, P. (2004). Italian and French teachers faced with pupil's academic failure: The norm of effort. *European Journal of Psychology of Education*, *19*, 147-166. doi: 10.1007/BF03173229

MATTEUCCI, M.C., GUGLIELMI, D. & LAUERMANN, F. (2017). Teachers' sense of responsibility for educational outcomes and its associations with teachers' instructional approaches and professional wellbeing. *Social Psychology of Education*. doi: 10.1007/s11218-017-9369-y

MATTEUCCI, M.C. & HELKER, K. (2018). Who is responsible for educational outcomes? Responsibility ascriptions for educational outcomes in a sample of Italian teachers, parents, and students. *Learning and Individual Differences*, *61*, 239-249. https://doi.org/10.1016/j.lindif.2017.12.009

MATTEUCCI, M.C., CARUGATI, F., SELLERI, P., MAZZONI, E. & TOMASETTO, C. (2008). Teachers' judgment from a European psychosocial perspective. In F. Columbus (Ed.), *Teachers and teaching: Strategies, innovations and problem solving*. New York NY: Nova Science Publishers.

MATTEUCCI, M.C., TOMASETTO C., SELLERI P. & CARUGATI F. (2008). Teacher judgments and pupils' causal explanations: Social valorization of effort-based explanations in school context. *European Journal of Psychology of Education*, *23*, 421-432.

McLACHLAN, G. & KRISHNAN, T. (2007). *The EM algorithm and extensions* (*Vol. 382*). Hoboken, NJ: John Wiley.

MEREDITH, W. & TERESI, J.A. (2006). An essay on measurement and factorial invariance. *Medical Care*, *44*, S69-S77. doi: 10.1097/01.mlr.0000245438.73837.89

RICHARDSON, P.W., KARABENICK, S.A. & WATT, H.M. (2014). *Teacher motivation: Theory and practice*. New York, NY: Routledge.

ROSE, J.S. & MEDWAY, F.J. (1981). Measurement of teachers' beliefs in their control over student outcome. *The Journal of Educational Research*, *74*, 185-190. doi: 10.1080/00220671.1981.10885308

SCHERMELLEH-ENGEL, K., MOOSBRUGGER, H. & MÜLLER, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, *8* (2), 23-74. Retrived from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.509.4258&rep=rep1&type=pdf

SCHNEIDER, S.L. (2013). The International Standard Classification of Education 2011. In G.E. Birkelund (Ed.), *Class and stratification analysis (Comparative Social Research, Volume 30*). Bingley, UK: Emerald Group Publishing Limited. doi: 10.1108/S0195-6310(2013)0000030017

SCHREIBER, J.B., NORA, A., STAGE, F.K., BARLOW, E.A. & KING, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. T*he Journal of Educational Research*, *99*, 323-338. doi: 10.3200/JOER.99.6.323-338

SHAVER, K.G. (1985). *The attribution of blame: Causality, responsibility, and blameworthiness*. New York, NY: Springer-Verlag.

SILVERMAN, S.K. (2010). What is diversity? An inquiry into preservice teacher beliefs. *American Educational Research Journal*, *47*, 292-329.

STEIGER, J.H. & LIND, J.C. (1980, May). *Statistically based tests for the number of common factors*. Paper presented at the Annual meeting of the Psychometric Society, Iowa City, IA.

TUCKER, L.R. & LEWIS, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*, 1-10. doi: 10.1007/BF02291170

VANDENBERG, R.J. & LANCE, C.E. (2000). A review and synthesis

of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*, 4-70. doi: 10.1177/109442810031002

WINTER, P.A., BRENNER, D.B. & PETROSKO, J.M. (2006). Teacher job satisfaction in a reform State: The influence of teacher characteristics, job dimensions, and psychological states. *Journal of School Leadership*, 16, 416-437.

WITT, L.A. (2001). Responsibility in work organizations. In A.E. Auhagen & H.W. Bierhoff (Eds.), *Responsibility: The many faces of a social phenomenon*. New York & London: Routledge.

WOSNITZA, M., HELKER, K. & LOHBECK, L. (2014). Teaching goals of early career university teachers in Germany. *International Journal of Educational Research*, *65*, 90-103.

APPENDIX

Teacher Responsibility Scale (TRS) and corresponding Italian version

Imagine that the following situations would occur when you have classes of your own. To what extent would you feel PERSONALLY responsible that you should have prevented each of the following?	Immagini che le seguenti situazioni si verificassero nella Sua classe. In che misura si sentirebbe PERSONALMENTE responsabile e/o si sentirebbe di aver dovuto impedire ciascuna delle seguenti situazioni?
I would feel PERSONALLY responsible if…	Mi sentirei PERSONALMENTE responsabile se …..
…a student of mine was not interested in the subject I teach (RSM1)	…un mio studente non fosse interessato alla materia che insegno (RSM1)
…a student of mine did not like the subject I teach (RSM2)	…un mio studente non amasse la materia che insegno (RSM2)
…a student of mine did not value learning the subject I teach (RSM3)	…un mio studente non considerasse importante l'apprendimento della materia che insegno (RSM3)
…a student of mine failed to make excellent progress throughout the school year (RSA1)	…un mio studente non riuscisse a fare eccellenti progressi durante l'anno scolastico (RSA1)
…a student of mine failed to learn the required material (RSA2)	...un mio studente non riuscisse ad imparare il materiale richiesto (RSA2)
…a student of mine had very low achievement (RSA3)	...un mio studente ottenesse risultati molto scarsi (RSA3)
…a student of mine thought he/she could not count on me when he/she needed help (RRS1)	…un mio studente avesse pensato di non poter contare su di me quando aveva bisogno di aiuto (RRS1)
…a student of mine did not think that he/she could trust me with his/her problems in or outside of school (RRS2)	…un mio studente pensasse di non potersi fidare di me se ha dei problemi all'interno o all'esterno della scuola (RRS2)
…a student of mine did not believe that I truly cared about him/her (RRS3)	...un mio studente non credesse che io veramente mi interesso a lui / lei (RRS3)
…a lesson I taught was not as effective for student learning as I could have possibly made it (RTE1)	…una lezione che ho svolto non fosse stata, per l'apprendimento degli studenti, così efficace quanto invece avrei potuto fare (RTE1)
…a lesson I taught was not as engaging for students as I could have possibly made it (RTE2)	…una lezione che ho svolto non fosse stata così coinvolgente come invece avrei potuto fare (RTE2)
…a lesson I taught failed to reflect my highest ability as a teacher (RTE3)	…una mia lezione non riflettesse le mie più alte capacità come insegnante (RTE3)

Note. All items were rated on a 7 point Likert scale from 0 = Not at all responsible to 6 = Completely responsible.
Legenda. RSM = Responsibility for Student Motivation; RSA = Responsibility for Student Achievement; RRS = Responsibility for Relationships with Students; RTE = Responsibility for Teaching.