

Datum

augustus 2005

© Hogeschool Utrecht,
Kenniskring Gedragsproblemen in de Onderwijspraktijk,
Utrecht, 2005

Bronvermelding is verplicht.
Vereenvoudigen voor eigen gebruik of intern
gebruik is toegestaan.

KG-publicatie nr. 1
Behaviorally Challenging
Students and Teacher Stress

H.A. Everaert

University of Professional Education of Utrecht

J.C. van der Wolf

University of Amsterdam and University of Professional Education of Utrecht

Paper presented at the annual meeting of the American Educational Research Association (AERA), Montréal, Canada, April 11-15, 2005

Correspondentie over deze KG-publicatie sturen naar:

huub.everaert@hvu.nl

KG-publicaties bevatten interne notities, verslagen en prepublicaties bestemd voor spreiding in kleine kring. De verantwoordelijkheid voor de inhoud van een KG-publicatie berust bij de auteur(s). Uit KG-publicaties mag alleen geciteerd worden met toestemming van de auteurs.

Voorwoord

De Kenniskring Gedragsproblemen in de Onderwijspraktijk aan de Hogeschool van Utrecht bestaat uit ongeveer tien personen met als lector J.C. van der Wolf.

Het lectoraat is opgericht in 2003 en heeft als taken het verrichten van vooral toegepast onderzoek en het helpen bij de ontwikkeling van curricula voor de onder de Faculteit Educatieve Opleidingen vallende afdelingen.

Door de 'leden van het lectoraat' worden allerlei stukken, artikelen, plannen, nota's etc. geproduceerd die dan in verschillende kringen worden verspreid zodat niet altijd een goed beeld ontstaat van de activiteiten van de Kenniskring. Via deze KG-publicatiereeks wil de kenniskring een vaste groep geïnteresseerden in staat stellen om op de hoogte te blijven van de activiteiten. Verwacht mag worden dat de KG-publicatiereeks uit zeer verschillende, uiteenlopende stukken zal gaan bestaan.

Voor U ligt de eerste KG-publicatie nr.1 van Everaert en Van der Wolf over 'Behaviorally Challenging Students and Teacher Stress'. U wordt van harte uitgenodigd om commentaar te leveren.

KG-publicaties

nr.1 Everaert, H.A. en J.C. van der Wolf (2005)
Behaviorally Challenging Students and Teacher Stress

Behaviorally Challenging Students and Teacher Stress

Abstract

The present study focuses on the level of stress a teacher perceives when dealing with the most behaviorally challenging student in his or her classroom. To measure stress in Dutch elementary classrooms, a sample was drawn of 582 teachers. Two questions concerning this relation between student and teacher will be addressed. First of all, we focus on background variables of teachers and students as sources of variation in explaining the magnitude of challenging student behavior and the associated level of stress teachers experience. The second topic of this paper is to accommodate the potentially stressful relationship between student and teacher in a wider network of surrounding variables, which are, Self-efficacy, Negative affect, Autonomy in taking decisions, and Support amongst colleagues. To evaluate the presence of challenging behavior, the behavior of the student is related to more general variables like student responsibility, class size and ratio of boys to girls. We close our paper by assessing the validity of the studied relationship between teacher and student with respect to possible burnout.

Keywords: behaviorally challenging students, teacher stress, efficacy, negative affect, autonomy, level of support, student responsibility, class size and ratio of boys to girls

Theoretical approach

Student behaviour stressors are one of the main sources of psychological distress among teachers (Borg, 1990). With the influx of exceptional students into regular classrooms, within the framework of world-wide inclusion and integration oriented policies, teachers in regular schools often incur new and additional duties for which they have either limited or no formal training (Brophy, 1996; Palmer Wilson, Gutkin, Hagen, & Oats, 1998). Due to the students' problematic behavior and lower abilities, the teachers also experience minimal and infrequent pupil progress (Coladarsi, 1992). Continual exposure to challenging behavior, both from pupils and their parents, can seriously deplete the teacher's emotional and physical resources, leading to self-doubt, loss of satisfaction from teaching, impulsivity, rigidity or feelings of anger and guilt (Coie & Koepl, 1990; Van der Wolf & Defares, 1994). In an attempt to provide the required services, work overload and hence stress are almost inevitable. In this respect, considerable strain has been placed upon the coping resources of teachers (Borg, 1990; Boyle, Borg, Falzon & Baglioni, 1995; Van der Wolf & Everaert, 2003). As a result teachers may react negatively and irritated to problem children. Children then may not receive the human contact, attention and support they need. This in turn can result in problem-behavior (Baker, 1999).

Much has been written about the origins and consequences of teacher self-efficacy (Tschannen-Moran, Woolfolk Hoy, & Hoy, 1998). Teacher efficacy has been defined as "the extent to which the teacher believes he or she has the capacity to affect student performance". (Bergman, McLaughlin, Bass, Pauly & Zellman, 1977, p. 137.) Self-efficacy affects the effort teachers invest in teaching, the goals they set, and their level of aspiration. Greater self-efficacy enables teachers to be less critical of students when they make errors (Ashton & Webb, 1986), to work

longer with a student who is struggling (Gibson & Dembo, 1984), and to be less inclined to refer a difficult student to special education (Soodak & Podell, 1993). Research on *efficacy perceptions* (Bandura, 1997) has shown that effort and persistence are greater in individuals who view themselves as competent or efficacious. They believe that they are capable of performing a task successfully and thus earning the rewards that success brings. High teacher self-efficacy in general has been linked to a variety of positive outcomes. Gibson and Dembo (1984) observed that teachers with high self-efficacy are more positive and responsive to students. They experience less stress (Greenwood, Olejnik, & Parkay, 1990). These results undoubtedly describe teachers who are more likely to promote positive classroom environments, which would result in better relationships with students.

Confrontation with deviant pupils often leads to a negative attitude, causing frustration because of the inability to give expression to these feelings. The basic cause of a negative attitude is that incidents in the classroom (unruly behaviour, immature behaviour or inadequate academic progress) become overwhelming. Incidents and conflict situations with a problem pupil that in turn negatively affects the atmosphere in a class cause feelings of anger and irritation in teachers. Naturally, this results in a negative, distancing attitude towards the problem pupil. Once this attitude has been established, the chances are good that every successive confrontation (and in primary education there are many of these confrontations) will cause stress and irritation (Lamude & Scudder, 1992; Lamude, Scudder & Furno-Lamude, 1992). So, teachers' stress is associated with interactional problems with students. Teachers are more likely to express personal feelings of anger toward disruptive students (Durivage, 1989). The negative patterns of interaction between teachers and students may contribute to increased conflicts and lack of understanding, thus leading to unsatisfactory relationships.

Apart from self-efficacy and negative affect, a lot has been written about teacher stress and the autonomy to take decisions. Teachers are often considered powerless. That is, teachers are perceived as subordinates at the bottom of the educational hierarchy (Datnow, 2000). The main problem here is the struggle teachers have to face when implementing curricular policies stipulated by others. This struggle is evident when teachers need to alter curricular platitudes to meet specific students' academic and behavioral needs (Reed, 2000; Taylor Webb, 2002). When teachers feel they have freedom in their department, and believe they are in control of their workload, generally the level of stress they experience is reduced. Allowing staff to take control of their workload mostly cause a reduction in stress levels. Teachers favour clearly defined roles, and within limits, support the idea of limited prescription, clear instruction and confident management (Bush, 1995). The level of support is another factor that affects a teacher's level of stress (Ursprung, 1986; Schaufeli, 1998). If the level of support from a head of school or a teacher's colleagues is low, teachers fragment into small groups and become rivals which can result in high level of stress and conflict.

It goes without saying that the relationship between challenging behavior and teacher stress also needs to be evaluated against the classroom atmosphere. The issue of student responsibility and self-discipline is of increasing interest in schools. In various countries, there is talk of the need for school curriculum to result in the production of more responsible citizens (Kennedy, 1996; Print, 1995; Anderson, Avery, Pederson, Smith & Sullivan, 1997; Lickona, 1996;

Bickmore, 1997;Osborne, 1995). The factors underlying such concern range from increasing student violence (Kauffman and Burbach,1997) to research which indicates that even students report that there are too many disruption in classrooms (Benninga and Wynne,1998).

Sociologists and educationalists have given a great deal of thought to what might be the reasons for boys being disadvantaged. Diefenbach and Klein (2002), for example, point in their highly respected study "Bringing boys back in" to a correlation between the overrepresentation of women in the teaching profession and the poorer performance of boys. Female teachers, they claim, are likely to value the behaviour of boys and girls differently.

"Female teachers dominate the school culture and possibly expect and reward the type of behaviour that girls are taught as part of their socialization process, and boys are not (to the same extent). In contrast, behavioral patterns which disrupt lessons and presumably also have a negative effect on performance in school are more commonly found in boys than girls, and female teachers may perhaps find this behaviour more annoying than male teachers if they are basing their standards on their own gender-specific socialization" (Diefenbach & Klein, 2002, p. 943).

Obviously, they go on to say, this theory would need to be tested by a relevant study.

Until recently, the classroom processes that differentiate small from large classes have proven remarkably intangible. In spite of the overabundance of measures, most of the findings were negative (Slavin, 1990). No differences were found in pupil satisfaction or affective measures, and no differences were found for most teacher activities, subject emphasis, classroom atmosphere (Shapson, Wright, Eason & Fitzgerald, 1980). When class size was changed, teachers did not alter the proportion of their time spent in interacting with the whole class, with groups, or with individual pupils. However, an opinion exists among parents and teachers that smaller classes are better than larger classes. Allan (1992) extended the starting-point of traditional class size research by studying teacher workload, feelings of stress and morale and the impact of student diversity. Teachers observed that students with Exceptional Educational Needs required extra attention and more preparation time if they are to be successful in the classroom. The STAR grade 4 follow-up demonstrated that students who had been in smaller classes were less disruptive than their peers in regular classes were. The Success Starts small project documented that grade 1 disciplinary referrals dropped over successive years in small classes. Further, Klein (1985) observed that even though class size did not change the degree of individualized instruction, the teacher did spend up to twice as much time per student in reduced size classes. In the present study we assess the effect of class size with respect to teacher's perception of the occurrence of problem behavior.

In the next paragraph we describe our sample and measures. Characteristics of students and teachers will be considered in explaining challenging behavior and teacher stress. A related question is to accommodate the potentially stressful relationship between student and teacher against a wider network of the above-mentioned surrounding variables. To assess validity of the studied relationship we will study the influence of the burnout model as presented by Byrne (1994).

Method

Sample and Procedures

Participants in this study were initially 607 regular primary teachers in the Netherlands. Prior to analysis, variables were examined on missing values. Of the total sample, 25 respondents did not fill out the questionnaire to the end and were deleted from the sample. Missing Value Analysis (MVA) did not result in significant differences ($\chi^2 = .05$) between the deleted respondents and the ones on which further analyses are based. The results in this paper are based on these 582 teachers (79% Female). Average years of teaching experience were 15 years. Participants were mostly recruited in the Dutch provinces of Noord-Holland, Zuid-Holland, Zeeland, Utrecht, Gelderland and Noord-Brabant.ⁱ The study was held in October-November 2004 with a response rate of 35%.

Measures

To measure the interaction between teacher and student, teachers had to think of the most behaviorally challenging student in the ongoing school year. A demographic questionnaire of the student teachers had in mind was administered. Questions dealt with the gender, age, ethnicity, DSM-IV diagnoses, SES and family situation of the student.

Inspired by Jere Brophy's *Teaching Problem Students* (1996) a first study was already made in autumn 2001 by the authors to measure problematic student behavior in the Netherlands (Everaert, 2003). Brophy (1996) distinguished four categories of twelve different types of behaviorally challenging pupils.ⁱⁱ We operationalized the vignettes describing problematic behavior as given by Brophy in 72 different items. As said, teachers had to think only of the most behaviorally challenging student while rating these items, an idea originally put forward by Greene, Abidin, and Kmetz (1997). Two samples were drawn in 2001. In the first sample teachers had to Q-sort 60 out of the original 72 items (N=122). In the other sample (N=154) teachers had to rate 72 items describing problematic student behavior. The scoring dimension in the second sample ranged on a 5-point Likert scale from 1 to 5. Based on these two samples, the conclusion was reached that using Likert response formats for scoring challenging student behavior resulted in reliable information (Everaert, 2003). Also, the results were used to improve the formulation of the respective items. Some items were rejected and others reformulated. In October 2003 another study was held to evaluate the usefulness and psychometric quality of the reformulated items. In total, 316 teachers in primary schools were asked to rate on a 5-point Likert scale both incidence and perceived stress caused by problematic behavior. Principal Axis Factoring was used to find the underlying structure and resulted in 38% of explained variance of incidence items. Principal Axis Factoring of the perceived stress items resulted in the same six factors with 46% of variance explained. Six different subscales were discerned: (1) *Against the grain*, (2) *Full of activity/Easily distractible*, (3) *Needs a lot of attention/Week student*, (4) *Easily upset*, (5) *Failure syndrome/Excessively perfectionistic*, and (6) *Aggressive/Hostile*.

Measuring challenging student behavior in this study is based on these former results. As might be clear by now, teachers were asked to think of the most behaviorally challenging student in their classroom and had to rate just 22 items to tap 6 different subscales for both incidence and perceived stress. The scoring dimension ranges on a 5-point Likert scale from 0 to 4.ⁱⁱⁱ Coefficient alpha reliability for the six scales measuring the incidence of problematic behavior ranged from .70 (Easily upset) to .81 (Needs a lot of attention/Week student). For scales

measuring perceived stress coefficient alpha was even higher and ranged from .72 (Failure syndrome/excessively perfectionistic) to .85 (Full of activity/Easily distractible). In Table 1, two exemplary items of every scale are presented.

Table 1 *Exemplary Items of Different Scales to Measure Incidence and Perceived Stress of Behaviorally Challenging Students (N = 582)*

Scale	Items
<i>Against the grain</i>	Seeks conflicts with teachers purposely Goes against the rules deliberately
Full of activity Easily distractible	Wanders around the classroom more than most others Much more active than others
Needs a lot of attention/ Weak student	Work always gives him or her a great deal of trouble Everything has to be predicted
Easily upset	Gets upset easily over smallest things Cries more often or has more stress than other pupils
Failure syndrome/ Excessively perfectionistic	Is generally not at all pleased with the final results Hands in work giving remarks such as; it will be wrong anyway
Aggressive/Hostile	Can be very destructive Often damages other pupils things

Whereas teachers had to think of one particular student while filling out items dealing with behaviorally challenging students, the remainder of the questionnaire dealt with general attitudes and notions about school climate. Seven items were included to measure *Self-efficacy in relationship building and behavioral management*. Using a 7-point scale from 1 (*not true at all*) to 7 (*very true*) participants responded to statements like “I can build a good relationship with even the most difficult student”, “I have positive characteristics that are very helpful when there is a problem with a student”, and “I can successfully handle the situation when one of the students gets disruptive or oppositional”. Cronbach’s alpha for the seven items of the Self-Efficacy scale was .80. The following three items were used to measure teacher’s *Negative affect*: “I have difficulty controlling my emotions when there is a conflict with students”, “I feel angry when a student repeatedly does not follow my advice”, and “students hurt my feelings by intentionally not following my directions”. Cronbach’s Alpha for this scale was relatively low with .61. Scales Self-efficacy and Negative affect were both developed by Yoon (2002).

Negative self-appraisal was measured using part of a scale developed by Carver & Ganellan (1983) and adapted by Wearing, Bell, McMurray, Conn, & Dudgeon (1990). The three-item scale measures the extent to which teachers are self-critical. Items included “when even one thing goes wrong I begin to feel bad and wonder if I can do well at anything at all”, “if something goes wrong – no matter what it is – I see myself negatively”, and “I often change from feeling extremely good about myself to seeing only the bad in me and feeling like a failure”. Teachers were asked to rate the extent to which these items described themselves on a five-point scale

ranging from 1 (*definitely untrue*) to 5 (*definitely true*). The internal homogeneity of Negative self-appraisal was .79.

In total, seven items like “the quality of your relationships with co-workers” and “the extent to which your co-workers stimulate you and support you in your work” were employed to tap *Support amongst colleagues*. A five-point Likert response format ranging from 1 (*very unsatisfied of this school characteristic*) to 5 (*very satisfied of this school characteristic*) was used to score these seven items making up scale Support amongst colleagues. Internal homogeneity of this scale was very satisfactorily with a score of .89.

To capture the teacher’s opinion of student behavior three different scales were used. Scale *Student responsibility and discipline* is characterised by five items like “the degree of responsibility students show toward their school assignments” and “your overall level of satisfaction with student responsibility and discipline in your school”. The scores on three items were administered to measure *Student behavioral values*. The last scale dealing with student behavior was made up of four items dealing with *Student-peer relations* and was constructed by using items like “Students care about each other” and “Students respect each other”. The scoring dimension of these twelve items ranged from 1 (very unsatisfied of this school characteristic) to 5 (very satisfied of this school characteristic). Indices of internal consistency ranged from .74 for Student behavioral values to .86 for Student responsibility and discipline. Cronbach’s alpha for Student-peer relations was also high with .85.

Five five items were used to measure the degree of *Autonomy in making decisions*. Items like “I have to ask my supervisor before I do almost anything” or “even small matters have to be referred to someone higher up for a final answer” can be considered exemplary for this four-point Likert response format ranging from 1 (*definitely true*) to 4 (*certainly untrue*). Coefficient alpha reliability of .77 was well above the minimum value of .70 as advocated by Nunnally (1978).

At last, *Burnout* was measured using the Dutch version of the Maslach Burnout Inventory for Teachers (UBOS-L). The questionnaire includes 22 items divided into three subscales: *Emotional exhaustion* (EE, 8 items), *Personalization* (DP, 7 items) and *Personal accomplishment* (PA, 7 items). The items were measured on a 7-point Likert scale, ranging from 0 (*never*) to 6 (*every day*). Scores on the scales are added separately. High scores on the scales EE and DP and low scores on the PA scale are indicative of burnout. Instances of items are: “I feel emotionally drained from my work” (EE), “I feel burned out from my work” (EE), “I’ve become more callous toward people since I took this job” (DP), “I feel students blame me for some of their problems” (DP), “I feel exhilarated after working closely with my recipients” (PA), and “I have accomplished many worthwhile things in this job” (PA). We found internal homogeneity of .88 (EE), .79 (DP), and .62 (PA) respectively.

From the above line-up of items and scales, it is clear we used different Likert dimensions. The reason for this is found in the possibility to compare our results with data gathered by other scholars; a matter endorsed by APA guidelines that “the essence of the scientific method involves observations that can be repeated and verified by others” (APA, 2001, p.348). No respondent made any comment that the questionnaire was confusing by using varying Likert scoring dimensions.

Results

As expected, 86% of the 582 teachers had a boy in mind when asked to think of a behaviorally challenging student. About 72% of the students are raised in a nuclear family, while 19% is a member of a one-parent household headed by the mother. With respect to ethnicity, a fourth of the students has at least one parent who is not originally born in the Netherlands. Measured on the first of January 2004, the average age of students in regular elementary schools was 8 years old. We also asked the teachers to indicate whether a psychiatrist or psychologist had diagnosed the pupil. This was the case for 17% of the students. Attention-Deficit/Hyperactivity Disorder and PDDNOS were the two most frequently mentioned DSM-IV-RT diagnoses.

Table 2 *Means, Standard Deviation, Alpha Reliabilities for Behaviorally Challenging Students (N = 582)*

Scale	Incidence				Perceived Stress			
	<i>M</i>	<i>SD</i>	<i>Al-pha reliabil-ities</i>	<i>Num-ber of items</i>	<i>M</i>	<i>SD</i>	<i>Al-pha reliabil-ities</i>	<i>Num-ber of items</i>
Against the grain	1.8	1.0	.79	4	1.5	1.0	.82	4
Full of activity/Easily distractible	2.7	0.9	.80	4	1.7	1.0	.85	4
Needs a lot of attention/Weak student	1.6	1.1	.81	4	0.9	0.8	.82	4
Easily upset	2.2	0.9	.70	4	1.3	0.8	.76	4
Failure syndrome/Excessively perfectionistic	1.1	1.0	.76	3	0.6	0.7	.72	3
Agressive/Hostile	1.1	1.1	.80	3	1.0	1.1	.82	3

In Table 2 means and standard errors of incidence and stress of challenging behavior are reported. On average, perceived stress associated with the displayed behavior of the student is lower than the incidence of the behavior itself. Most stress is generated by Full of activity/Easily distractible students (1.7), closely followed by Against the grain (1.5). While aggressive/hostile behavior does not occur that often, it must be noted that the level of stress is almost just as high (1.0), as the incidence (1.1). It may be concluded that the perceived stress is highly dependent on the kind of challenging behavior itself. Pearson *r* product-moment correlations between the incidence and stress of challenging behavior are high and vary from .62 (Failure syndrome / Excessively perfectionistic) to .83 (Aggressive/Hostile).

First of all, we ran twelve ANOVA-models to examine whether child characteristics, which are gender, ethnicity, family situation at home, and DSM-IV-RT diagnoses might explain variance in challenging behavior and variance in related stress. Because of highly inflated Type I errors when employing several models, only models with a *p*-value <.005 were taken into further consideration.^{iv} Only with respect to Failure syndrome/Excessively perfectionistic children it turned out that two out of four child characteristics did matter (i.e., family situation at home and

DSM-IV-RT). However, straightforward conclusions with respect to perceived stress associated with Failure syndrome/excessively perfectionistic were hindered because the null hypothesis stating error variance of the dependent variable is equal across groups, was rejected ($F(25,508) = 1.57, p < .05$). Therefore, in the remaining of this study child characteristics are being ignored. Also from this point of view, there seems to be no logical reason against putting together the six discerned types of challenging student behavior as one overall scale.

Our next step was to study whether teacher gender played a role in singling out different types of problematic behavior (cf. Diefenbach & Klein, 2002). Again, several between-subjects analysis of variances (ANOVA) were performed on different subscales. In total, twelve models were run with challenging student behavior (six times) and perceived stress (six times) as dependent variables. Unfortunately, it was not opportune running ANCOVA models that included years of teaching experience as covariate. The assumption of homogeneity of slopes was severely and repetitively violated. Pearson r product-moment correlations between the different subscales measuring challenging behavior and perceived stress on the one hand and years of teaching experience was rather low. It varied between $-.07$ (Against the grain) and $-.15$ (Perceived stress of Full of activity/Easily distractible). With respect to challenging student behavior and perceived stress, female teachers score higher than their male counterparts. That is, female teachers report more challenging behavior and more stress as a result of this behavior. With a maximum p -value of $.005$ for model significance, means of female teachers for the incidence of challenging behavior differ significantly from the means of male teacher with respect to Against the grain, Full of activity/Easily distractible, Easily upset, and Aggressive/hostile. The results are printed in Table 3.

Table 3 Means for Behaviorally Challenging Students by Teacher Gender ($N = 582$)

Scale	Incidence		Perceived Stress	
	Female	Male	Female	Male
Against the grain ^a	1.9	1.6	1.5	1.3
Full of activity/Easily distractible ^a	2.8	2.4	1.8	1.5
Needs a lot of attention/ Weak student	1.6	1.4	0.9	0.8
Easily upset ^a	2.2	1.9	1.3	1.1
Failure syndrome/ Excessively perfectionistic	1.2	1.0	0.6	0.5
Aggressive/Hostile ^{a,b}	1.2	0.8	1.1	0.7

^a Significant model with respect to incidence ($p < .005$).

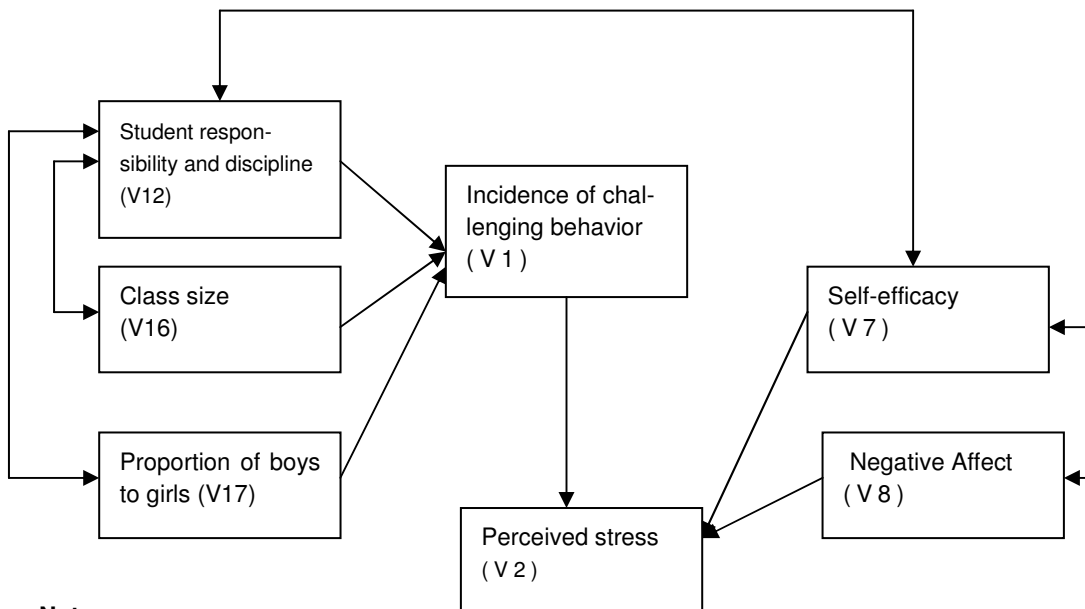
^b Significant model with respect to perceived stress ($p < .005$).

We used EQS-software (EQS 6.1) to run two different SEM models. In the first model, we tried to explain the quality of the teacher student-relationship in a wider network of school-related variables as described at the beginning of this paper. In the second model, we investigate whether Maslach's model of burnout can be used as a predictor of the potentially stressful relation between teacher and student. Different indices are recommended for evaluating the quality of SEM models (Byrne, 1994; Hu & Bentler, 1999; Mueller & Hancock, 2004; Ullman, 2001). We decided to report five different fit indices: the χ^2 statistic, the comparative fit index (CFI), the normed fit index (NFI), the non-normed fit index (NNFI) and the root mean-square

error of approximation (RMSEA). In general, the target values for the selected fit indices CFI, NFI and NNFI should be $\geq .95$, while a RSMEA $\leq .05$ indicates also good model fit. Good models produce usually consistent results on many different indices.

We started the first model with quite a number of explaining variables. The heart of the model is the causal relationship between teacher stress and a behaviorally challenging student. Based on the correlation matrix three variables were brought in to exert direct influence on the incidence of challenging behavior (i.e., Student responsibility and discipline, Number of children in classroom, and Proportion of boys in classroom). Apart from the incidence of challenging behavior, variables that were used to explain perceived stress included Self-efficacy, Negative affect, Negative self-appraisal, Relationship with colleagues, Support amongst colleagues, and Autonomy in making decisions. Also several correlation parameters to estimate covariance between the independent had to be estimated in fitting the model to the data. Unfortunately, the used model fit indices all indicate that the model does not accurately describe the sampled date, to say the least ($\chi^2_{(35)} = 85.326$, $p\text{-value} < .000$, CFI = .958, NFI = .932, NNFI = .934, and RMSEA = .052). The multivariate Wald test results suggested dropping Negative self-appraisal, Relationship with colleagues, Support amongst colleagues, and Autonomy in making decisions. After this was done, we got a clear-cut model ($\chi^2_{(11)} = 16.543$, $p\text{-value} = .1221$, CFI = .991, NFI = .973, NNFI = .982, and RMSEA = .031), see Figure 1. No further relations between variables were required or necessary on basis of the Lagrange Multiplier Test.

Figure 1 *Postulated Structural Model for Incidence of Challenging Behavior and Perceived Stress (N=582).*



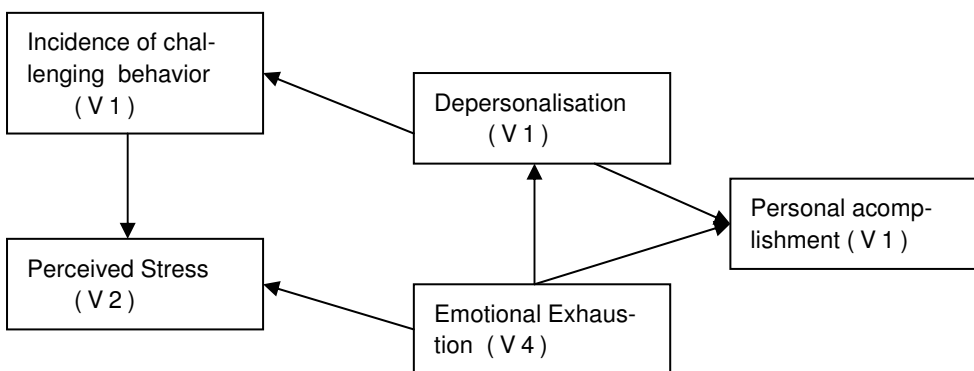
Note.

$V1 = -.198*V12 + .106*V16 + .102*V17 + .970 E1$.
 $V2 = .704*V1 - .114*V7 + .191*V8 + .660 E2$.
 $Cov(V12,V7) = .071$, $Cov(V12,V16) = .484$, $Cov(V12,V17) = -.007$, and $Cov(V7,V8) = -.357$.
 All reported standardized parameters are significant at the 5% level.

So, the relationship between a behaviorally challenging student and teacher stress is not to be understood in terms of collegial support and autonomy in making decisions as is generally advocated in stress theory. Also the influence of Negative self-appraisal can be discarded. What is indeed important in understanding stress is feeling competent in dealing with students, like Self-efficacy, also the influence of Negative affect is considerably. We are fully aware that the topic of model-fit evaluation is very complex. The proposed model tried to incorporate relevant theoretical knowledge, and to underline our approach, it is better to speak of a postulated model. Also, the validity of the postulated model should be evaluated in replication studies.

In the second structural equation model we look at the question whether the relation between the incidence and perceived stress is related to burnout. Where most scholars treat burnout as a dependent variable, in the second model the three related scales (i.e., Emotional exhaustion, Depersonalisation, and Personal accomplishment) indicative of burnout are used to assess the validity of the relation between incidence and stress. The specification how to relate Emotional exhaustion, Depersonalisation, and Personal accomplishment is given by Byrne (1994). As stated in the first paragraph, substantive theory suggests that Depersonalization will have an effect on the incidence of challenging behavior. At the same time, emotional exhaustion and perceived stress are positively related. We would like to repeat that burnout has to be considered as a continuing process over time, involving all sources of stress, including the behavior of several students. In the proposed model, we combine this general outline with the behavior of just one student viewed by the teacher. The model is fitted to the data and as expected, burnout scales are indicative of the proposed relation between the incidence and stress as a result of the interaction between student and teacher. In figure 2, the fitted model is presented ($\chi^2_{(4)} = 5.329$, p-value = .2552, CFI = .998, NFI = .992, NNFI = .995, and RMSEA = .025).

Figure 2 *Postulated Structural Model for Incidence of Challenging Behavior, Perceived Stress, and Burnout (N=582).*



$$V1 = .180 \cdot V5 + .984 E1; \quad V2 = .698 \cdot V1 + .165 \cdot V4 + .683 E2.$$

$$V5 = .469 \cdot V4 + .883 E5. \quad V6 = -.236 \cdot V5 - .237 \cdot V4 + .914 E6.$$

No covariances were specified. All reported standardized parameters are significant at the 5% level.

Discussion

In general, sources of variance with respect to incidence and perceived stress of challenging student behavior could not be explained by referring to child characteristics. The same may be concluded with respect to teacher characteristics, except that female teachers notify more behavioral problems and experience more stress than their male counterparts when asked to pick out the most behaviorally student. In fact, this may be seen as an indication of the cited theory of Diefenbach and Klein (2002) where they illicit the troubles of boys dealing with growing numbers of female teachers in primary education. However, to underline their point of view, interaction effects between gender of teachers and students should be taken into account. In another paper we address this topic in detail.

Once again, we have to state explicitly that cited theories explaining educational stress do not focus on one teacher dealing with one child, as we did in this study. We cannot, nor do we want to, reject the general influence of variables like autonomy, support of colleagues, or negative self-appraisal on the basis of this study, but we can claim that in understanding the relationship between teacher and student Negative affect and Self-efficacy are important. However, more substantial research is necessary to validate our model. On the other hand, given the role of Negative affect and Self-efficacy in dealing with challenging student behavior, they offer practical and relatively easy ways of training teachers how to deal with behavioral problems. Anyway, that will be a lot easier than changing the behavior of students. Especially, if one realizes that there will always be a second 'best' behaviorally challenging student to confront.

References

- Allan, R. (1992). *Class Size and Teaching Assignments of WEAC's Teacher Members*. Madison (Win.): Wisconsin Education Association Council.
- American Psychological Association. (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington, DC: American Psychological Association.
- Anderson, C. Avery, P. G. Pederson, P. V. Smith, E. S. & Sullivan, J. L. (1997) Divergent perspectives on citizenship education: A Q-method study and survey of social studies teachers. *American Educational Research Journal*, 34(2), 333-64.
- Ashton, P.T., & Webb, R.B. (1986). *Making a difference: Teachers' sense of efficacy and student achievement*. New York: Longman.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Baker, J.A. (1999). Teacher student interaction in urban at-risk classrooms: Differential behavior, relationship quality, and student satisfaction with school. *Elementary School Journal*, 100(1), 57-70.
- Benninga, J. S. & Wynne, E. A. (1998) Keeping in Character. *Phi Delta Kappan*. Feb 1998: 439-445.
- Bergman, P. , McLaughlin, M. , Bass, M., Pauly, E., & Zellman, G. (1977). *Federal programs supporting educational change: Vol. VII. Factors affecting implementation and continuation*. Santan Monica, CA: Rand.
- Bickmore, K. (1997). *Teaching conflict and conflict resolution in school:(Extra-) curricular considerations*. Paper presented at Connections 1997 International Social Studies Conference. July. Australia.
- Borg, M.G. (1990). Occupational stress in British educational settings: a review. *Educational Psychology*, 10, 103-126.
- Boyle, G.J., Borg, M.G., Falzon, J.M., & Baglioni, A.J. (1995). A structural model of the dimensions of teacher stress. *British Journal of Educational Psychology*. 65(1), 49-67.
- Brophy, J. (1996). *Teaching problem students*. New York: The Guilford press.
- Bush, T. (1995). *Theories of Educational Management*. London: Paul Chapman.
- Byrne, B.M. (1994). *Structural Equation Modeling with EQS and EQS/Windows. Basic Concepts, Applications, and Programming*. Thousand Oaks: Sage Publications.
- Carver, C.S. & Ganellan, R.J. (1983). Depression and components of self-punitiveness: high standards, self-criticism, and overgeneralization. *Journal of abnormal Psychology*, 92, 330-337.
- Coie, J.D., & Koepl, G.K. (1990). Adapting intervention to the problems of aggressive and disruptive children. In S. R. Asher & J.D. Coie (Eds.). *Peer rejection in childhood* (pp. 309-337). New York, NY: Cambridge University Press.
- Coladarci, T. (1992). Teachers' sense of efficacy and commitment to teaching. *Journal of Experimental Education*, 60, 323-337.
- Coladarci, T. (1992). Teachers' sense of efficacy nad commitment to teaching. *Journal of Experimental Education*, 60, 323-337.
- Datnow, A. (2000). Power and Politics in the Adoption of school Reform Models. *Educational Evaluation and Policy Analysis*, 22, 357-374.
- Diefenbach, H. and Klein, M., Bringing Boys back in. *Soziale Ungleichheit zwischen den Geschlechtern im Bildungssystem zuungunsten von Jungen am Beispiel der Sekundarabschlüsse*. [Bringing boys back in. Social inequality between the sexes in the education system to the detriment of boys, using the example of secondary qualifications], in: Zeitschrift für Pädagogik, 2002,48, 938-958.
- Durivage, A. (1989). Assaultive behaviour: Before it happens. *Canadian Journal of Psychiatry*, 34, 393-397.

- Everaert, H.A. (2003). Het meten van de meester. [Measuring Teacher Stress]. In: Wolf, K. van der, (Ed.). *Het hoofd van de meester*. Utrecht: Uitgeverij Agiel.
- Everaert, H.A. & Van der Wolf, K.C. (2004). *Stress in the Student-Teacher Relationship in Dutch Schools: a replication study of Greene, Abidin & Kmetz's Index of Teaching Stress-Instrument (ITS)*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA, April 12-16.
- Greene, R.W., Abidin, R.R., Kmetz, C. (1997). The Index of Teaching Stress: A measure of student-teacher compatibility. *Journal of School Psychology, 35*(3), 239-259.
- Gibson, S. & Dembo, M.H. (1984). Teacher Efficacy: A construct validation. *Journal of Educational Psychology, 74*, 4, 569-582.
- Greenwood, G.E., Olejnik, S.F., & Parkay, F.W. (1990). Relationships between four teacher efficacy belief patterns and selected teacher characteristics. *Journal of Research & Development in Education, 23*, 2, 102-106.
- Hu, L. & Bentler, P.M. (1999). Cutoff Criteria for Fit Indices in Covariance Structure Analysis: Conventional Criteria versus New Alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*, 1-55.
- Kauffman, J. M. & Burbach, H. J. (1997) On creating a climate of classroom civility. *Phi Delta Kappan*, Dec: 320-325
- Kennedy, K. J. (1996). New Challenges for Civics and Citizenship. ACT.Australia ACSA.
- Klein, K.R. (1985). Practical Applications of Research: Class Size. *Phi Delta Kappan, 66*, 578-580.
- Lamude, K.G. & Scudder, J. (1992). Resistance in the College Classroom: Variations in Students' Perceived Strategies For Resistance and Teachers' Stressors as a Function of Students Ethnicity. *Perceptual and Motor Skills, 75*, 2, 615-626.
- Lamude, Kevin G., Scudder, Joseph, & Furno-Lamude, Diane (1992). "The Relationship of Student Resistance Strategies in the Classroom to Teacher Burnout and Teacher Type-A Behavior," *Journal of Social Behavior and Personality, 7*, 597-610.
- Lickona, T. (1996). Eleven Principles of Effective Character Education. *The Journal of Moral Education, 25*, 1, 93-100.
- Mueller, R.O. & Hancock, G.R. (2004). *Evaluating Studies that Utilize Structural Equation Modeling: Guidelines for Manuscript Reviewers*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA, April 12-16.
- Nunnally, J.C. (1978). *Psychometric Theory*. New York: McGraw-Hill.
- Osborne, K. (1995). In Defence of History teaching in the Past and the Meaning of Democratic Citizenship. Toronto. Our School-Our Selves Educational Foundation.
- Print, M (Ed). (1995). Civics and Citizenship: Issues from Practice and Research. ACT. Australia. ACSA.
- Palmer Wilson, C., Gutkin, T.B., Hagen, K.M., & Oats, R.G. (1998). General education teachers' knowledge and self-reported use of classroom interventions for working with difficult-to-teach student: Implications for consultation, preferral intervention and inclusive services. *School Psychology Quarterly, 13*(1), 45-62.
- Reed, C. (2000). Teaching with Power: shared decision-making and classroom practice. New York: Teachers College Press.
- Schaufeli, W.B. (1998) *Burnout and Lack of Reciprocity Among Teachers*. Paper presented at a Symposium on Teacher Burnout held at the Annual Meeting of the American psychological association, San Francisco, CA, August.
- Shapson, S.M., Wright, E.N., Eason, G., & FitzGerald, J. (1980). An Experimental Study of Class Size. *American Educational Research Journal, 17*, 2, 141.

- Slavin, R.E. (1990). Class Size and Student Achievement: Is Smaller Better? *Contemporary Education*, 62,6.
- Soodak, L. & Podell, D. (1993). Teacher Efficacy and student problems as factors in special education referral. *Journal of Special Education*, 27, 66-81.
- Taylor Webb, P. (2002). Teacher Power: the exercise of professional autonomy in an era of strict accountability. *Teacher Development*, 6, 1, 47-61.
- Tschannen-Moran, M., Woolfolk Hoy, A., & Hoy, W.K. (1998). Teacher Efficacy: Its meaning and measure. *Review of Educational Research*, 68, 202-248.
- Ullman, J.B. (2001). Structural Equation Modeling. In Tabachnick, B.G. & Fidell L.S. (Eds.), *Using Multivariate Statistics* (pp. 653-658). Boston: Allyn and Bacon.
- Ursprung, A.W. (1986) Burnout in the human services: a review of the literature. *Rehabilitation Counseling Bulletin*, 29, 190-199.
- Wearing, A.J., Bell, R., McMurray, N.E., Conn, M. & Dudgeon, P. (1990). *Teacher Stress in Victoria: A Survey of Teachers' Views. The Report*. Melbourne: Education Shop.
- Wolf, J.C. van der & Defares P.B. (1994). Stress bij leraren in het basisonderwijs [Stress in primary school teachers] In: N. Ph Geelkerken, G.J.J. Goetheer, F.H.J.G. Brekelmans & F.A.J. van Moorsel, red., *Praktijkboek bedrijfsgezondheidszorg in het onderwijs. Deel C: Instrumenten, hulpmiddelen en methoden*.(pp. 45-53) Den Haag: VUGA.
- Wolf, K. van der & Everaert, H. (2003). Teacher Stress, Challenging Parents and Problem Students. In S. Castelli, M. Mendel, & B. Ravn (Eds.), *School, Family, and Community Partnership in a World of Differences and Changes* (pp. 135-146). Gdansk: Wydawnictwo Uniwersytetu Gdanskiego.
- Yoon, J.S. (2002). Teacher characteristics as predictors of teacher-student relationships: stress, negative affect, and self-efficacy. *Social Behavior and personality*, 30 (5), 485-494.

Notes

-
- ⁱ Several students of the University of Professional Education of Utrecht participated in collecting the data: Arjanneke Brandsma, Sabine Bax, Menno van Es, Petra den Hollander, Frits van het Hout, Gea Hoving, Gerbert Sipman, Lindy Slingerland, Albert Sluiter, Ingrid Muurman, Gerda Pool en Wil Vlam. We appreciate their efforts in sampling the respondents.
- ⁱⁱ The four categories described by Brophy (1996) are students with achievement problems, students with hostility problems, student role-adjustment problems, and students with social relationship problems.
- ⁱⁱⁱ Greene, Abidin, and Kmetz (1997) formulated the shining idea to think only of the most behaviorally challenging student while rating the items. We feel indebted to them. The reason to change the scoring dimension from 1 to 5 into 0 to 4 is based on another study held also held in October 2003 (Everaert and Van der Wolf, 2004).
- ^{iv} Because $1-(.995)^{12} \approx .95$, we choose p-values $\leq .005$ as limit.