Measuring general and specific stress causes and stress responses among beginning secondary school teachers in the Netherlands.

Harmsen, Ruth; Helms-Lorenz, Michelle; Maulana, Ridwan; van Veen, Klaas; van Veldhoven, Marc

Published in:
International Journal of Research & Method in Education

DOI:
10.1080/1743727X.2018.1462313

Publication date:
2019

International Journal of Research & Method in Education

ISSN: 1743-727X (Print) 1743-7288 (Online) Journal homepage: https://www.tandfonline.com/loi/cwse20

Measuring general and specific stress causes and stress responses among beginning secondary school teachers in the Netherlands

Ruth Harmsen, Michelle Helms-Lorenz, Ridwan Maulana, Klaas van Veen & Marc van Veldhoven

To cite this article: Ruth Harmsen, Michelle Helms-Lorenz, Ridwan Maulana, Klaas van Veen & Marc van Veldhoven (2019) Measuring general and specific stress causes and stress responses among beginning secondary school teachers in the Netherlands, International Journal of Research & Method in Education, 42:1, 91-108, DOI: 10.1080/1743727X.2018.1462313

To link to this article: https://doi.org/10.1080/1743727X.2018.1462313

Published online: 20 Apr 2018.

INTERNATIONAL JOURNAL OF RESEARCH & METHOD IN EDUCATION
2019, VOL. 42, NO. 1, 91–108
https://doi.org/10.1080/1743727X.2018.1462313

Routledge
Taylor & Francis Group

Measuring general and specific stress causes and stress responses among beginning secondary school teachers in the Netherlands

Ruth Harmsen [a], Michelle Helms-Lorenz[a], Ridwan Maulana[a], Klaas van Veen[a] and Marc van Veldhoven[b]

[a]Teacher Education, University of Groningen, Groningen, the Netherlands; [b]Human Resource Studies, Tilburg University, Tilburg, the Netherlands

ABSTRACT

The main aim of this study was to adjust the Questionnaire on the Experience and Evaluation of Work (QEEW) in order to measure stress causes and stress responses of beginning secondary school teachers in the Netherlands. First, the suitability of the original QEEW stress scales for use in the beginning teachers (BTs) context was investigated using a sample of 356 beginning teachers from 52 different secondary school locations in the Netherlands. Confirmatory Factor Analyses, Principal Component Analyses and Mokken scaling item reduction was applied to create high concise and precise scales. Hereafter, based on the teacher stress literature, additional teacher specific stress items were added, resulting in the adjusted version of the measure, the Questionnaire on the Experience and Evaluation of Work – Beginning Teachers (QEEW-BT, study 1). To cross-validate the results and to examine the internal consistency and validity of the adjusted instrument a different sample of 143 beginning teachers from 61 different secondary school locations in the Netherlands was used (study 2). The present findings provide adequate support that the QEEW-BT is a reliable and valid instrument to measure stress causes and responses for beginning secondary school teachers in the Netherlands.

ARTICLE HISTORY
Received 11 May 2017
Accepted 10 January 2018

KEYWORDS
Stress; questionnaire; beginning teachers; secondary education

Introduction

Knowledge about work stress, its causes and negative consequences, and how it affects employees' wellbeing and performance has been well established (Van Veldhoven 1996; Van Veldhoven and Meijman 1994). General stress causes – as measured by the Questionnaire on the Experience and Evaluation of Work (QEEW, in Dutch: VBBA) of Van Veldhoven and Meijman (1994) – are positively related to stress responses, and negatively related to wellbeing and performance. The QEEW has robust psychometric properties and has been widely used and recognized in the Netherlands and various other countries (e.g. Belgium, France, Germany, Italy, and Brasil). Despite its strong psychometric quality and popularity, the QEEW is limited for use in a job-specific context such as the teaching profession. The questionnaire does not cover specific stress factors related to the teaching context. Although we assume that general stress causes and responses apply to teachers' work as well, inclusion of stress factors specifically related to the teaching context is necessary to unravel the knowledge about stress in the teaching profession more comprehensively. In the present study,

CONTACT Ruth Harmsen r.harmsen@rug.nl Teacher Education, University of Groningen, Grote Kruisstraat 2/1, Groningen 9712 TS, the Netherlands

we aim to adjust the QEEW to measure stress causes and stress responses among beginning teachers (BTs).

The teaching profession is considered to be a highly stressful profession by nature (Johnson et al. 2005). In the Netherlands, in 2014, circa 1 in 5 teachers experienced burnout symptoms. Teachers also reported higher levels of workload compared to other professionals (Hooftman et al. 2015).

There is no consistent definition of (teacher) stress provided by the experts in the field and competing bodies of literature on teacher stress exist that originate from different fields (e.g. Izawa et al. 2012; Klassen and Chiu 2010). That being said, there seem to be two general perspectives on (teacher) stress. The first one is that stress responses (e.g. tension) are a result of something ouside of the individual, external factors (e.g. heavy workload). For example Kyriacou (2001) and Rudow (1999) define teachers stress as teachers' experience of unpleasant, negative emotions, such as tension and anger, resulting from some aspects of their work. The other perspective posits that stress is internal; it is what goes on inside the individual as they interpret or react to what is going on around them (Gold and Roth 1993). For example the Transactional model of Stress (Folkman 2013; Lazarus and Folkman 1984) views work stress as a result of an interaction and appraisal process between the employee and its environment. In the same line, other researchers conceptualize stress with both internal and external aspects: the degree of mismatch between the demands made upon an individual and the individual's ability to cope with those demands (Bakker and Demerouti 2007; McCarthy et al. 2016). Given the on-going debate about and the different usages of the term 'teacher stress' in this paper the term is used more as a label indicating a specific field of (applied) research. We aim to develop an instrument that can be useful in both externally and internally focussed teacher stress research.

In this study teacher stress is devided into: (1) stress causes, (2) stress responses, and (3) stress outcomes. Stress causes are the collection of aspects of the work content and the work situation that influence employees at a cognitive, motivational and emotional level, for example student misbehaviour. Stress responses are the employees' mental interpretation when experiencing stress causes, for example feeling tension (Van Veldhoven 1996). Stress outcomes result from persistent stress causes and responses, for example leaving the teaching profession (attrition). To understand the chain between beginning teachers' stress causes, responses, teaching behaviour and attrition the well-validated and widely used job demands-resources (JD-R) model is used (Bakker and Demerouti 2007). According to the JD-R model, there are two main psychological processes at work. The first is called the health impairment process. This process describes the relationship between job demands, job resources, strain and organizational outcomes. The second proces is motivational in nature. Herby it is assumed that job resources have motivational potential and lead to high work engagement, low cynicism, and excellent performance. The focus in this study is to adjust an instrument to measure the outcomes of the health impariment proces. The JD-R model posists that job demands cover 'physical, psychological, social, or organizational aspects of the job that require sustained physical and/or psychological (cognitive and emotional) effort or skills'. Examples in the teaching context are many hours of actual teaching and student misbehaviour. Although job demands are not necessarily negative, they may turn into stress causes when meeting those demands requires high effort from which the employee has not adequately recovered. Hence, job demands are conceptually related to stress causes. Job resources refer to physical, psychological, social or organizational aspects of the job that are either/or: functional in achieving work goals, reduce job demands and the associated physiological and psychological costs, stimulate personal growth, learning and development (Bakker and Demerouti 2007). According to the JD-R model: high job demands exhaust employees' mental and physical resources leading to strain. Similarly, in the teaching context stress causes can lead to among other things feelings of tension. Therefore, strain is conceptually related to stress responses. Finally, this strain can lead to negative organizational outcomes like poor performance, health-related problems, and absenteeism (Bakker, Demerouti, and Sanz-Vergel 2014). Translated to the teaching context, tension can eventually lead to negative stress outcomes. Examples of negative

stress outcomes are leaving the teaching profession (dropout) and less effective teaching behaviour. Effective teaching behaviour refers to teachers' behaviour which has significant and positive impact on student learning and outcomes (Maulana, Helms-Lorenz, and van de Grift 2015). According to Van de Grift (2007, 2014), effective teaching behaviours which are observable from the teacher's work include: creating a safe and stimulating learning climate, efficient classroom management, clear instruction, activating learning, teaching learning strategies and adaptation.

Stress causes that seem to recur consistently in most of the teacher stress literature involve education specific workload (e.g. big classes), student misbehaviour, poor relationships at work (with students, supervisor, colleagues), role ambiguity, lack of job autonomy, and poor school ethos (Borg and Riding 1991; Chen and Miller 1997; Hanif 2004). Therefore, it is important that these teaching-related stress causes are included in measuring teacher stress. These stress causes, except for the role in the organization, are found to be positively related to dropout (Buchanan 2010; Gonzalez, Brown, and Slate 2008; Shen 1997; Struyven and Vanthournout 2014; Wilhelm, Dewhurst-Savellis, and Parker 2000). Furthermore, greater levels of stress that are caused by student misbehaviour, education specific workload, and poor relationships at work have a negative impact on teachers' perceived personal accomplishments (Burke and Greenglass 1993; Kokkinos 2007). Hanif, Tariq, and Nadeem (2011) studied the relationship between stressors and teaching performance of teachers with a minimum one year of teaching experience. The study showed that stress experienced by teachers was negatively related to their teaching behaviour.

BTs are more vulnerable to the pressures of the profession and stress compared to more experienced teachers (Gold and Roth 1993). This group also shows high dropout rates (Macdonald 1999). Further, BTs show less effective teaching behaviour compared to more experienced teachers (Maulana, Helms-Lorenz, and van de Grift 2015). Dropout is harmful for student achievement (Ronfeldt, Loeb, and Wyckoff 2013) and less effective teaching behaviour can also influence student achievement negatively (Hattie 2012).

Insight into which and how stress causes and responses influence BTs' dropout and teaching behaviour remains unclear. This insight could help us signal the possibility of dropout at an early stage in the career, and to develop support which potentially decreases the negative influences of certain stress causes on teachers' teaching behaviour. To this end a valid, reliable and apt instrument to measure teacher stress more comprehensively is necessary.

As mentioned above, although there are several (teacher) stress questionnaires, none of them cover the broad scope of stress causes and stress responses. The Stressor Multilevel context scale of Betoret (2006) and the Stress questionnaire of Payne and Furnham (1987) are both teacher stress questionnaires which do not cover teacher specific stress causes such as poor relationship with students. Additionaly, those questionnaires do not include stress responses factors, which limit the insight into the mental interpretation of teachers when experiencing stress causes. The Teacher Stress Inventory of Fimian (1984) lacks the important teacher stress causes such as student misbehaviour, role ambiguity, and poor relationship with students, supervior, and colleagues. In addition, the teacher stress questionnaire of Kyriacou and Sutcliffe (1978) and the stress questionnaire of Borg and Riding (1991) do not include two of the important stress causes which seem to recur consistelty in most of the teacher stress literature, namely poor relationship with students and role ambiguity.

None of the mentioned teacher stress questionnaires cover the broad scope of stress causes and stress responses that we are aiming for. Therefore, the knowledge base on teacher stress will benefit from the construction of a more comprehensive teacher stress measure covering the broad scope of stress causes and stress responses experienced by BTs. Because the QEEW has been proven to be a robust and widely used questionnaire measuring general stress across professions including teaching, the QEEW will be used as the point of departure to create the Questionnaire on the Experience and Evaluation of Work – Beginning Teachers (QEEW-BT), and we will be both modifying the source and adding scales to it to make it more specific for the context of BTs.

The studies

Two studies were conducted using different samples of secondary education teachers in the Netherlands. Study 1 focuses on adjusting the QEEW to measure general and teacher specific stress causes and stress responses of beginning teachers (BTs), resulting in the concept version of the QEEW-BT. Study 2 focuses on cross-validating the factor structure of the QEEW-BT and on examining the internal consistency, scalability, construct validity and criterion validity of the QEEW-BT, resulting in the final version of the QEEW-BT. Permission to conduct the studies was taken from the concerned authorities of all the schools before the studies were conducted. Also, participants were aware that participation was voluntary and that they could stop participating at any stage of the studies.

Study 1: adjusting the QEEW resulting in the QEEW-BT

Method

Participants

Sample 1 consisted of 356 beginning teachers from 52 different secondary school locations in the Netherlands (see Table 1). The percentage of female teachers is slightly higher (56.7%) compared to the national secondary school teachers population. The percentage of school locations with less than 1000 students is lower than in the national population. The percentage of school locations with between 1000 and 2000 is higher in this sample than in the national population. The percentage of qualified teachers is lower than in the national population, both compared to teachers from all ages as well as teachers who are younger than 35 years. The distribution of the denomination, urbanization and SES percentages differ greatly between the national population and the sample. This can be explained by the distribution of the school locations over the regions in the Netherlands (see Table 2). The northern regions of the Netherlands (Drenthe, Friesland, and Groningen) are overrepresented. In these regions there are only two catholic school locations. Also, the urbanization and the SES in these regions is lower compared to most of the other regions in the Netherlands.

Measures

BTs from sample 1 completed the QEEW around the end of the school year 2012 (April, measurement occasion 1). BTs were defined as teachers who recently obtained their teaching qualification or would obtain this within a year and had less than three years of teaching experience. In addition, one class of students per BT filled in the student questionnaire measuring perceived teaching behaviour (students' perceptions of the teachers' ability on the six domains described below; Maulana, Helms-Lorenz, and van de Grift 2015) also around April for three consecutive years (measurement occasion 1, 2, 3). For each school the contact person of the school randomly chose one of the BTs' classes to fill in the student questionnaire, and a survey administrator supervised the students whilst they were completing the questionnaire (the BT was not allowed to be present). Students were instructed and informed that participation was voluntary and that their answers would be treated anonymously.

Perceived teaching behaviour. Perceived teaching behaviour was measured using an unidimensional student questionnaire consisting of the six domains; safe and stimulating learning climate, efficient classroom management, clear instruction, activating learning, differentiated teaching, and teaching-learning strategies. A total of 6291 students completed the student questionnaire for 281 of the BTs from sample 1 on measurement occasion 1, and a total of 4811 students for 208 of the BTs from sample on measurement occasion 3. The individual student ratings were aggregated to classroom level for which mean averages were used. The instrument had a good internal consistency (Cronbach's alpha = .80), and consisted of 24 questions rated on a four-point Likert scale ranging from 1 (completely disagree) to 4 (completely agree) (Maulana, Helms-Lorenz, and van de Grift 2015). Change in perceived teaching behaviour was operationalized as the gain score between

Table 1. Information on the samples and population.

		All secondary schools in the Netherlands	Sample 1	Sample 2
Number of teachers		73900[a]	356	143
Percentage female		46.7%[a]	56.7%	57.3%
Schools[b]		655	25	25
School locations[b]		1432	52	61
Number of students per school location	less than 1000 students	70.5%	65.4%	67.2%
	between 1000 and 2000 students	24.9%	32.7%	31.1%
	between 2000 and 3000 students	1.7%	0%	0%
	missing	2.9%	1.9%	1.6%
Teaching qualification	qualified[d]	88.2%[c]	63.8%	95.8%
	appointable[e]	6.2%[c]	3.9%	0.7%
	not qualified[f]	5.6%[c]	19.4%	0.7%
Teaching qualification people < 35 years	qualified	81.1%[c]	63.8%	95.8%
	appointable	8.7%[c]	3.9%	0.7%
	not qualified	10.2%[c]	19.4%	0.7%
	missing		12.9%	2.8%
Denomination	public (in the Netherlands: openbare scholen)	27.2%	51.9%	49.2%
	catholic	21.2%	1.9%	1.6%
	protestan/christian/evangelical/reformed	33.7%	34.6%	37.7%
	free schools (in the Netherlands: Algemeen bijzondere scholen)	16.1%	5.8%	6.6%
	remaining	1.8%	3.8%	3.3%
	missing		2.0%	1.6%
Urbanization	1 (most)	22.4%	9.6%	9.8%
	2	31.4%	11.5%	9.8%
	3	21.1%	34.6%	34.4%
	4	17.7%	23.1%	21.3%
	5 (least)	6.2%	19.2%	21.3%
	missing	1.2%	2.0%	3.3%
Social economic status (SES)	4st (lowest)	32.35	43.42	45.36
	3st	23.61	28.25	26.64
	2st	19.75	14.78	14.03
	1st (highest)	23.52	11.23	11.54

[a]Data from 2013. Source: OCW (Ministerie van Onderwijs, Cultuur en Wetenschap 2014).
[b]Data from 2014. Source: data from Dienst Uitvoering Onderwijs (DUO).
[c]Data from 2014. Source: IPTO bevoegdheden en vakken in het VO (Fontein et al. 2016).
[d]Teachers who obtained a teaching degree who teach the same subject and education level they have been qualified for.
[e]In the Netherlands teachers are considered to be appointable if they comply to become qualified within two years after being appointed. This category of teachers consist of different types of teachers: teachers with no teaching degree who teach no longer than one year. Teachers who teach a different subject or education level to what they have been qualified for. Pre-service teachers. Second career teachers who are not qualified yet.
[f]Teachers with no teaching degree who teach longer than one year and are not following a teacher education programme.

Table 2. Distribution of school locations (correspond to Dutch provinces) in both samples (in percentages).

Province	School locations (%)	Sample 1	Sample 2
Drenthe	3.1	23.1	23.0
Flevoland	2.7	1.9	1.6
Friesland	5.7	30.8	31.1
Gelderland	11.7	7.7	6.6
Groningen	4.8	28.8	31.1
Limburg	5.2		
Noord-Brabant	12.7		
Noord-Holland	15.4		
Overijssel	7.8	5.8	4.9
Utrecht	6.3		
Zeeland	2.0		
Zuid-Holland	22.6		
Total	100	100	100

measurement occasion 1 and 3. For 176 of the BTs from sample 1 the student questionnaire was completed both on occasion 1 and occasion 3, representing 4580 student questionnaires.

Stress causes and stress responses. Stress causes and stress responses were measured with the original QEEW (Van Veldhoven and Meijman 1994). All 27 scales (201 items) were used (see Appendix A for an overview of the scales). The internal consistency and validity of this questionnaire was good (Evers, Van Vliet-Mulder, and Groot 2000). Most scales had a good internal consistency (Cronbach's alpha > .80) and five scales had a satisfactory internal consistency (Cronbach's alpha between .65 and .79) in our sample.

Dropout. Dropout was registered between measurement occasion 1 and 3 when teachers left their first teaching job (1 = job leaver, 0 = stayer).

Analytic approach

In order to adjust the QEEW to measure general and teacher specific stress causes and stress responses of beginning secondary school teachers in the Netherlands, four systematic steps were followed.

Step 1. Examining and selecting the QEEW scales which are applicable for the beginning teacher context. The QEEW has 27 scales of which we intend to select only those scales that capture the broader scope of stress causes and stress responses. This was achieved by examining and selecting the scales from the QEEW which belonged to second order stress causes and stress responses scales. Second order scales are aggregated scales that measure more aspects (derived from the primary scales) of one latent construct. To examine whether the second order factor structures for the stress causes and stress responses of BTs are similar to that of employees from a variety of occupations as indicated by Van Veldhoven (Van Veldhoven, Appendix B structure 1), Confirmatory Factor Analyses (CFAs) were conducted at the scale level (see Appendix C for an explanation of the technical and statistical terms). The second order factor structures of BTs appeared to be different from that of employees from a variety of occupations. Therefore, the factor structures were explored by means of Principal Component Analyses (PCAs). The PCAs reflect a formative measurement model, meaning that the measured variables are considered to be the cause of the latent variable. In Appendix B both second order factor structures can be found (structure 1 and 2a and 2b). The new factor structure still consisted of 4 factors representing stress causes (structure 2a) and 2 factors representing stress responses (structure 2b). However, role conflict which belonged to the factor of social and organizational aspects in the structure for employees from a variety of occupations belongs to the factor of high psychological task demands in the BT structure. Lack of job autonomy belonged to lack of influence, whereas in the BT structure it belonged to the factor lack of development opportunities. Lack of communication which originally belonged to the factor social and organizational aspects appeared to belong to the factor lack of influence in the BT structure. Finally, poor sleeping quality and tiredness during work belong to the stress response tension in the BT structure. Scales within the PCA factor structure (all scales in structure 2a and 2b) were selected for our purposes to capture the broader scope of stress causes and responses.

Step 2. This step was used to identify scales having significant predictive values for dropout or change in perceived teaching behaviour using regression analyses. Scales that revealed predictive value were selected for our purposes.

Step 3. In order to create a more cost-effective teacher stress instrument with high conciseness and precision (thus avoiding redundancy) and increased user friendliness (i.e. reducing fatigue effects) item reduction was needed. Items from the scales selected in step 1 and 2 were deleted stepwise during three rounds of item reduction, which was consistent with the procedure taken by earlier initiatives for item reduction of the QEEW (Notelaers et al. 2007; Van Veldhoven et al. 2015). Consistent with the procedure taken by Van Veldhoven (1996), Mokken scaling item reduction procedure

was conducted. More specifically, the more restrictive double monotonicity Mokken model was applied (Mokken 1971). The assumption of Unidimensionality was tested by conducting the Mokken analysis for each stress scale. Items with item-scalability values (H(i)-values) below .30 violated the assumption of Unidimensionality and were therefore deleted. The assumption of local independency, was tested by means of the LD X^2 statistic (Chen and Thissen 1997). Standardized X^2 values above 10 indicate violation of the local independency assumption. The assumption of non-intersecting item response curves was checked using the information from non-intersection based on P-Matrix. *Crit* value above 80 is a strong violation of the assumption of non-intersecting response curves (Molenaar and Sijtsma 2000).

In the second round of item reduction, items measuring the same content were deleted. This was achieved by checking the mean score of the items. Items with the same mean score (or within a distance of .10) were checked in terms of their content. If the items were similar in content the item with the lowest H(i) score was deleted. One exception to this rule was made for scales with a significant predictive value for change in perceived teaching behaviour and/or dropout. Items in those scales were only deleted if the predictive value of the scale for perceived teaching behaviour and/or dropout remained significant after deletion.

The last round of deletion was content-based. The aim was to reduce the amount of original items in the scale to about 50%. Factor loadings of the items on the scale were calculated and the items with the lowest factor loading were deleted stepwise if after deletion the scale internal consistency was minimal .70, the scalability (H(t)) minimal .40 and the amount of items per subtopic in the scale was well balanced. That is, the scale lack of job variety for example has items regarding 'Task variety' (subject A) and items regarding 'Creativity' (subject B). The reduced scales has three items regarding subject A and three regarding subject B.

Step 4. As the QEEW measures general stress causes and stress responses, during this step teacher stress causes based on the teacher stress literature were added.

An overview of the steps can be found in Table 3 below.

Results

Step 1: second order factor structure of the QEEW

For the stress causes the CFA-model showed poor fit ($\chi^2 = 297.32$, $df = 59$, $p < .001$; RMSEA = .11; CFI = .81; TLI = .75; SRMR = .08). The CFA-model for stress responses also showed poor fit ($\chi^2 = 92.33$, $df = 8$, $p < .001$; RMSEA = .18; CFI = .86; TLI = .73; SRMR = .07). In both models the modification indices indicated ways to improve model-data fit by allowing correlations between scales from the same factor. Nevertheless, allowing scales to correlate did not result in acceptable model-data fit.

Table 3. Overview of the steps.

Step	Aim	Reason	Analyses	Results
1	Examining the factor structure of the QEEW for beginning teachers.	To identify and select scales of the QEEW contributing to a higher order scale. In order to capture the broader scope of stress causes and stress responses.	CFA, PCA	See step 1: Second order factor structure QEEW
2	Examining which of the scales which were not selected in step 1 have a predictive value for dropout and/or teaching behaviour.	To include scales of the QEEW which do not belong to a higher order scale that do have predictive value for dropout and/or teaching behaviour.	Regression analyses	See step 2: Predictive value
3	Item reduction on the scales which were selected during step 1 and 2.	To make the questionnaire more concise and precise. Hereby reducing the amount of items teachers have to fill in.	Mokken	See step 3: Item reduction
4	Adding teacher specific scales.	To make the questionnaire suitable for the teaching context.		See step 4: Additional scales

PCA was conducted on the 19 stress causes scales using (in line with the development of the original QEEW) orthogonal rotation (varimax). Preliminary analyses supported the adequacy to conduct PCAs. Initial analyses were run to obtain eigenvalues for each component in the data. Six components had eigenvalues over Kaiser's criterion of 1 and in combination explained 61.63% of the variance. Given the relatively large sample size, and the convergence of the scree plot and Kaiser's criterion on six components, four components were retained in the final analysis. PCA was repeated this time with four fixed factors (see Appendix B structure 2a for the factor structure and the factor loadings).

PCA was also conducted on the eight stress responses scales. Two components had eigenvalues over Kaiser's criterion of 1 and in combination explained 62.79% of the variance (see Appendix B structure 2b, and Appendix A for an overview of which stress causes and responses scales were selected in step 1).

Step 2: predictive value

None of the non-selected scales (scales which were not part of a higher order scale, step 1) had significant correlations with change in perceived teaching behaviour. Therefore no regression analysis was performed and no scales were selected based on this criterion.

Two scales, uncertainty about the future and changes in tasks, showed significant correlations with dropout. Therefore, logistic regression analysis was conducted to predict dropout using these two scales. Uncertainty about the future was a significant predictor of dropout ($\chi^2 = 11.675$, $p < .001$, $df = 1$ *with Exp(b)* $= 1.61$, $b = .48$, $SE = .14$, Nagelkerke $R^2 = .06$). Therefore, this scale was retained (see Appendix A, step 2).

A total of 22 scales had a predictive value for dropout and/or were part of a higher order scale and were therefore selected for the QEEW-BT. The five scales physical effort, lack of possibilities for contact, lack of career opportunities, lack of remuneration, and changes in tasks were not included in the QEEW-BT (25 items). These scales were not part of a higher order scale (step 1) and did not predict dropout and/or teaching behaviour (no evidence of predictive validity).

Step 3: item reduction

Eight items were deleted as they violated the assumption of Unidimensionality. Four item pairs violated the assumption of local independency. For three out of these four item pairs, the content of the two questions was clearly different, therefore no items were removed. For the last item pair the content was similar. Therefore, the item with the lowest H(i) score was removed. Lastly, two items were deleted as they violated the assumption of non-intersecting response curves.

In the second round of item deletion a total of 20 items were deleted. In the final round of item deletion another 16 items were deleted. A total of 129 items from the QEEW were used for the QEEW-BT.

Step 4: additional scales

Based on the literature review, additional stress causes which need to be included in a teacher stress questionnaire involve: education specific workload (e.g. big classes), student misbehaviour, poor relationships at work (students, supervisor, colleagues), role ambiguity, lack of job autonomy, and poor school ethos. Stress causes concerning poor relationships with supervisor, colleagues, role ambiguity and lack of job autonomy were already included in the QEEW. The other stress causes were not included in the original QEEW and were therefore added to the QEEW-BT (46 items, see Table 4 for an overview of the additional scales, example of items and the sources).

The items regarding poor relationship with students were divided into two scales: poor relationship with students 1 and poor relationship with students 2. Both scales cover items measuring teacher-student relationships. However, both scales differ in terms of the response category. The

Table 4. Overview of new scales and sources.

Scale	Example items	Sources
Student misbehaviour	Students who are rude	Sources of Stress Questionnaire (Borg and Riding 1991), Teacher Stress Questionnaire (Kyriacou and Sutcliffe 1978), Stress Questionnaire (Payne and Furnham 1987).
Poor school climate	experiencing lack of recognition for the work you do	
Poor relationship with students	Finding it difficult to like your students	Teacher as a Social Context Questionnaire (Wellborn et al. 1992), subscales affection and attunement. Student-Teacher Relationship Scale (Koomen et al. 2012), subscale closeness and conflict.
Education specific workload	Too many hours of actual teaching	Teacher Stress Inventory (Fimian 1984).

items of the first scale were provided on a four-point Likert scale ranging from 0 (completely disagree) to 3 (completely agree), while those of the second scale were scored on a five-point Likert scale ranging from 1 (definitely not applicable) to 5 (certainly applicable).

An overview of the steps, scales and number of items per scales can be found in Appendix A.

Study 2: cross-validation

Method

Participants

Study 2 included a total sample of 143 beginning teachers from 61 different secondary school locations in the Netherlands (see Table 1). The percentage of female teachers is slightly higher (57.3%) compared to the national secondary school teachers population. The percentage of school locations with less than 1000 students is lower than in the national population. The percentage of school locations with between 1000 and 2000 students is higher in the samples than in the national population. The percentage of qualified teachers is higher than in the national population. The distribution of the denomination, urbanization and SES percentages differ greatly between the national population and the sample. Similar to sample 1, the northern regions of the Netherlands (Drenthe, Friesland, and Groningen) are overrepresented.

Measures

Stress causes and stress responses. Stress causes and stress responses were measured with the QEEW-BT (see Appendix A). The teachers received the questionnaire digitally and they had to fill it in within a three month period.

Perceived teaching behaviour. Perceived teaching behaviour was measured during the same period as the stress causes and responses. The instrument and procedure for this measure are similar to study 1. A total of 1736 students completed the questionnaire for 86 of the BTs of sample 2.

Analytic approach

Using sample 2 the scalability of the 46 newly constructed items were tested using Mokken scaling. The same assumptions and criteria were employed consistent with study 1. The internal consistency of all QEEW-BT scales was investigated by calculating the Cronbach's alpha and the H(t). The construct validity was investigated by correlating the new stress causes scales with the stress responses scales. The criterion validity was established by correlating the new teacher specific scales with perceived teaching behaviour, see Figure 1 for the path of influence. Finally, the factor structure of the QEEW-BT was explored using PCAs.

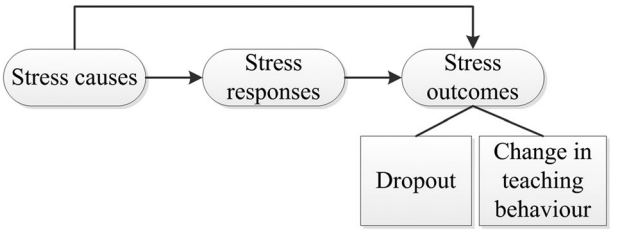

Figure 1. Path of influence stress causes, stress responses, and stress outcomes.

Results

Scalability

From the 46 newly constructed teacher specific items 5 items violated the assumption of unidimensionality and were therefore deleted. An overview of the final version of the QEEW-BT can be found in Appendix A.

Internal consistency

Appendix A indicates that all newly developed scales show satisfactory to good internal consistencies (Cronbach's alpha .76–.89). The QEEW-BT has 14 scales with good internal consistency and 13 scales with satisfactory internal consistency (Cronbach's alpha .70–.94). Further, all scales have sufficient H(t) values (.40–.82).

Construct validity

In Table 5 the results of the correlations between the new stress causes scales and the stress responses scales are displayed. The results show that the teacher specific stress causes have significant positive relationships with many of the stress responses (Spearman's rho ranging from .15 and .52, non-normally distributed data).

Criterion validity

In Table 6 the results of the correlations between the new stress causes scales with perceived teaching behaviour scale are displayed. The scales student misbehaviour, poor relationship with students 1

Table 5. Correlation (Spearman's rho) between new stress causes scales and stress responses scales.

Scale	Lack of work pleasure	Lack of organizational commitment	Turnover	Emotional reactions during work	Tiredness during work	Need for recovery	Poor sleep quality	Rumination
Student misbehaviour	.35**	.37**	.28**	.15*	.21**	.29**	.18*	.20*
Poor relationship with students 1	.39**	.44**	.22**	.20**	.20**	.28**	.18*	.07
Poor relationship with students 2	.45**	.29**	.07	.42**	.12	.20**	.19*	.04
Poor school climate	.10	.31**	.28**	.01	.32**	.36**	.17*	.35*
Education specific workload	.09	.22**	.27**	−.05	.40**	.52**	.33**	.50**

*Correlation is significant at the .05 level (1-tailed).
**Correlation is significant at the .01 level (1-tailed).

Table 6. Correlation (Spearman's rho) between new stress causes scales and perceived teaching behaviour.

Scale	Spearman's rho
Student misbehaviour	−.52*
Poor relationship with students 1	−.31*
Poor relationship with students 2	−.38*
Poor school climate	−.02
Education specific workload	−.03

*Correlation is significant at the .01 level (1-tailed).

and poor relationship with students 2 are negatively related to the perceived teaching behaviour scale (Spearman's rho −.52, −.31, and −.38).

Second order factor structure QEEW-BT

PCA was conducted on the 19 stress causes scales with orthogonal rotation (varimax). Initial analyses showed that the scales lack of communication and uncertainty about the future correlated with no other scales and were therefore not included in further analyses of the structure. The PCA was conducted on the remaining 17 stress causes scales with orthogonal rotation. Preliminary analyses supported the adequacy to conduct PCAs and were run to obtain eigenvalues for each component in the data. Five components had eigenvalues over Kaiser's criterion of 1 and in combination explained 66.06% of the variance (see Appendix B structure 3a for the factor structure and the factor loadings).

PCA was also conducted on the eight stress responses scales. Three components had eigenvalues over Kaiser's criterion of 1 and in combination explained 70.68% of the variance (see Appendix B structure 3b).

Conclusions and discussion

The main aim of this study was to adapt the Questionnaire on the Experience and Evaluation of Work (QEEW) in order to measure stress causes and stress responses of beginning secondary school teachers in the Netherlands. We found that stress of novice employees in the teaching professions manifests itself in a unique way. Suitable QEEW scales were selected and an item reduction procedure was applied. Furthermore, relevant teacher specific stress scales were added resulting in the QEEW-BT. In the second study the factor structure, internal consistency and validity of the QEEW-BT were examined, resulting in the final version of the QEEW-BT.

It was found that the majority of the original QEEW scales were relevant for use in the BTs context. The scales physical effort, lack of possibilities for contact, lack of career opportunities, lack of remuneration, and changes in tasks were evaluated to be not relevant. Those scales did not belong to a higher order scale and did not have predictive value for dropout and/or teaching behaviour. For the relevant scales the item reduction procedure resulted in a reduction to nearly half of the original questionnaire. For physical effort and changes in tasks a restriction of range seems evident. Most BTs (around 90%) in sample 1 never or only sometimes experienced stress caused by physical effort or changes in tasks. There was more variation on the scales lack of possibilities for contact, lack of career opportunities, and lack of remuneration. However, these stress causes did not seem to influence BTs dropout or perceived teaching behaviour.

Based on teacher stress literature, items regarding student misbehaviour, poor relationship with students, poor school climate and education specific workload were identified as relevant and added to the QEEW resulting in the QEEW-BT (see Apendix A for the final version). These additional scales had significant positive correlations with the stress responses scales, meaning that higher levels on one or more of the teacher specific stress causes scales were associated with higher levels on one or more of the stress responses scales. This is in line with findings of Mearns and

Cain (2003), adding to the construct validity of the scale and highlighting the importance of measuring these teacher specific stress causes.

More unique to the QEEW-BT, compared to other (teacher) stress questionnaires, are the scales regarding poor relationship with students. This study showed that poor relationships with students are positively related to lack of work pleasure, lack of organizational commitment, turnover, emotional reactions during work, tiredness during work, need for recovery, and poor sleep quality. Furthermore, a negative relationship between poor relationships with students and BTs perceived teaching behaviour was revealed. Therefore, there is evidence that poor relationships between the teacher and their students do not only influence students' school engagement and achievement negatively (Roorda et al. 2011), but also affect teachers' stress responses and perceived teaching behaviour in a negative fashion.

Student misbehaviour also showed a negative relationship with BTs perceived teaching behaviour. This is in line with earlier research showing that teachers with great levels of stress caused by student misbehaviour have lower levels of perceived personal accomplishment (Abel and Sewell 1999; Kokkinos 2007).

Most scales of the QEEW-BT had good internal consistencies and some had satisfactory internal consistencies. The PCA conducted on the stress causes of the QEEW-BT showed that the new scales education specific workload and poor school climate loaded on the existing higher order factors. Whereas, the scales poor relationship with students 1, poor relationship with students 2, and student misbehaviour together loaded on a new higher order factor, students. Thereby adding a new and important factor to the questionnaire. The PCA conducted on the stress responses also showed a new factor, negative emotions, with the scales emotional reactions during work and lack of work pleasure. These scales correlated highly with each other (Spearman's rho .66 in sample 2), meaning that higher levels on emotional reactions during work are associated with more lack of work pleasure. This is in line with the findings of Brackett et al. (2010), who found that the ability to regulate emotions is positively associated with job satisfaction in secondary school teachers.

It is interesting that the scales from the higher order scale students all have positive relationships with lack of work pleasure and emotional reactions during work, whereas the other new scales do not. It seems that students play an important role in the emotional reactions and work pleasure of BTs. This is in line with research of Fokkens-Bruinsma and Canrinus (2014) who found that for preservice secondary school teachers, working with children/adolescents is in the top three of most important motives to become a teacher. This motive has, in turn, a significant relationship with satisfaction, happiness and affective commitment.

Notwithstanding the strengths, the present study also has limitations. Although the samples were relatively large, both samples included mainly school locations in the northern regions of the Netherlands. Therefore, the distribution of the denomination, urbanization, and SES percentages in our samples differed to some extent from the national population profile. Hence, caution should be taken when interpreting results of this study until replication with larger and more representative samples is available. Another limitation was the size of the second sample ($N = 143$). Although this sample is sufficient, a larger sample offers more possibility to examine the relationships between the teacher specific scales and dropout, and change in perceived teaching behaviour with more power. Nevertheless, the QEEW was proven to be a robust questionnaire and has been used in many countries with very diverse cultural backgrounds (e.g. UK, USA, China, Malaysia, Australia, Japan, Brazil, Sweden). Also, both the items from the QEEW as well as from the QEEW-BT reflect the stress factors which are highlighted as important stress factors in international research. Another limitation of this study is the use of PCAs to establish the factor structure of the QEEW-BT. Although Mokken scaling was used next to PCA to determine the unidimensionality, the complicated issue of establishing the factor structure would in hindsight be better employed using another advance statistical method. Future research would benefit from using exploratory structural equation modelling (ESEM; Marsh et al. 2014) as an alternative for analysing stress factors at the item-level.

To conclude, the QEEW-BT offers an instrument that may improve our understanding of relevant BTs' work outcomes related to stress, which supports Sparks and Cooper's (1999) line of thinking advocating the combination of measuring general and specific stressors when assessing specific job settings. The adjustment of the QEEW has resulted in a more comprehensive measure to capture both general and specific stress causes among BTs – the QEEW-BT. Researchers and schools (i.e. coaches, mentors) can use this instrument to provide insight into which stress causes and responses BTs at their school experience in order to adjust their support more right on target to ensure the wellbeing of their (new) teaching staffs. Future research could focus on the question how these stress causes and stress responses influence BTs dropout and change in teaching behaviour.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by Ministerie van Onderwijs, Cultuur en Wetenschap [grant number OND/ODB-13/19888].

ORCID

Ruth Harmsen http://orcid.org/0000-0001-6567-6341

References

Abel, M. H., and J. Sewell. 1999. "Stress and Burnout in Rural and Urban Secondary School Teachers." *The Journal of Educational Research* 92 (5): 287–293.

Bakker, A. B., and E. Demerouti. 2007. "The Job Demands-resources Model: State of the Art." *Journal of Managerial Psychology* 22: 309–328.

Bakker, A. B., E. Demerouti, and A. I. Sanz-Vergel. 2014. "Burnout and Work Engagement: The JD–R Approach." *Annual Review of Organizational Psychology and Organizational Behavior* 1: 389–411.

Bandalos, D. L., and S. J. Finney. 2010. "Factor Analysis: Exploratory and Confirmatory." In *The Reviewer's Guide to Quantitative Methods in the Social Sciences*, edited by G. R. Hancock and R. O. Mueller, 93–114. New York, NY: Routledge.

Betoret, F. D. 2006. "Stressors, Self-efficacy, Coping Resources, and Burnout Among Secondary School Teachers in Spain." *Educational Psychology* 26 (4): 519–539.

Borg, M. G., and R. J. Riding. 1991. "Towards a Model for the Determinants of Occupational Stress Among Schoolteachers." *European Journal of Psychology of Education* 6 (4): 355–373.

Brackett, M. A., R. Palomera, J. Mojsa-Kaja, M. R. Reyes, and P. Salovey. 2010. "Emotion-regulation Ability, Burnout, and Job Satisfaction Among British Secondary-school Teachers." *Psychology in the Schools* 47 (4): 406–417.

Buchanan, J. 2010. "May I be Excused? Why Teachers Leave the Profession." *Asia Pacific Journal of Education* 30 (2): 199–211.

Burke, R. J., and E. Greenglass. 1993. "Work Stress, Role Conflict, Social Support, and Psychological Burnout Among Teachers." *Psychological Reports* 73 (2): 371–380.

Chen, M., and G. Miller. 1997. "Teacher Stress: A Review of the International Literature." Research report No. ED410187. http://files.eric.ed.gov/fulltext/ED410187.pdf.

Chen, W. H., and D. Thissen. 1997. "Local Dependence Indexes for Item Pairs Using Item Response Theory." *Journal of Educational and Behavioral Statistics* 22 (3): 265–289.

Cortina, J. M. 1993. "What is Coefficient Alpha? An Examination of Theory and Applications." *Journal of Applied Psychology* 78 (1): 98–104.

Evers, A., J. C. Van Vliet-Mulder, and C. J. Groot. 2000. "Documentatie van tests en testresearch in Nederland [Documentation of Tests and Test Research in the Netherlands]." In *Testbeschrijvingen*, 649–650. Assen: Van Gorcum & Comp. B.V.

Field, A. 2013. "Exploratory Factor Analysis." In *Discovering Statistics Using IBM SPSS Statistics*, edited by A. Field, 627–685. London: Sage.

Fimian, M. J. 1984. "The Development of an Instrument to Measure Occupational Stress in Teachers: The Teacher Stress Inventory." *Journal of Occupational Psychology* 57 (4): 277–293.
Fokkens-Bruinsma, M., and E. T. Canrinus. 2014. "Motivation for Becoming a Teacher and Engagement with the Profession: Evidence from Different Contexts." *International Journal of Educational Research* 65: 65–74.
Folkman, S. 2013. "Stress: Appraisal and Coping." In *Encyclopedia of Behavioral Medicine*, edited by M. D. Gellman, 1913–1915. New York, NY: Springer.
Fontein, P., P. Prüfer, K. de Vos, and A. Vloet. 2016. "IPTO: bevoegdheden en vakken in het vo [IPTO: Teaching Qualifications and Subjects in Secondary Education]." https://www.rijksoverheid.nl/documenten/rapporten/2016/06/17/ipto-bevoegdheden-en-vakken-in-het-vo.
Gold, Y., and R. A. Roth. 1993. *Teachers Managing Stress and Preventing Burnout: The Professional Health Solution*. London: The Falmer Press.
Gonzalez, L., M. S. Brown, and J. R. Slate. 2008. "Teachers who Left the Teaching Profession: A Qualitative Understanding." *The Qualitative Report* 13 (1): 1–11.
Hanif, R. 2004. "Teacher Stress, Job Performance and Self-efficacy of Women School Teachers." Doctoral thesis, National Institute of Psychology, Quaid-e-Azam University, Islamabad, Pakistan.
Hanif, R., S. Tariq, and M. Nadeem. 2011. "Personal and Job Related Predictors of Teacher Stress and job Performance among School Teachers." *Pakistan Journal of Commerce and Social Sciences* 5 (2): 319–329.
Hattie, J. 2012. *Visible Learning for Teachers: Maximizing Impact on Learning*. London: Routledge.
Hooftman, W. E., G. M. J. Mars, B. Janssen, E. M. M. de Vroome, and S. N. J. Van den Bossche. 2015. *Nationale Enquête Arbeidsomstandigheden 2014. Methodologie en globale resultaten* [National Working Conditions Survey 2014. Methodology and Overall Results]. Leiden: TNO.
Izawa, S., K. Saito, K. Shirotsuki, N. Sugaya, and S. Nomura. 2012. "Effects of Prolonged Stress on Salivary Cortisol and Dehydroepiandrosterone: A Study of a two-Week Teaching Practice." *Psychoneuroendocrinology* 37 (6): 852–858.
Johnson, S., C. Cooper, S. Cartwright, I. Donald, P. Taylor, and C. Millet. 2005. "The Experience of Work-related Stress Across Occupations." *Journal of Managerial Psychology* 20 (2): 178–187.
Klassen, R. M., and M. M. Chiu. 2010. "Effects on Teachers' Self-efficacy and Job Satisfaction: Teacher Gender, Years of Experience, and Job Stress." *Journal of Educational Psychology* 102 (3): 741–756.
Kokkinos, C. M. 2007. "Job Stressors, Personality and Burnout in Primary School Teachers." *British Journal of Educational Psychology* 77 (1): 229–243.
Koomen, H. M., K. Verschueren, E. Van Schooten, S. Jak, and R. C. Pianta. 2012. "Validating the Student-Teacher Relationship Scale: Testing Factor Structure and Measurement Invariance Across Child Gender and age in a Dutch Sample." *Journal of School Psychology* 50 (2): 215–234.
Kyriacou, C. 2001. "Teacher Stress: Directions for Future Research." *Educational Review* 53 (1): 27–35.
Kyriacou, C., and J. Sutcliffe. 1978. "Teacher Stress: Prevalence, Sources, and Symptoms." *British Journal of Educational Psychology* 48 (2): 159–167.
Lazarus, R. S., and S. Folkman. 1984. *Stress, Appraisal, and Coping*. New York, NY: Springer.
Macdonald, D. 1999. "Teacher Dropout: A Review of Literature." *Teaching and Teacher Education* 15 (8): 835–848.
Marsh, H. W., A. J. Morin, P. D. Parker, and G. Kaur. 2014. "Exploratory Structural Equation Modeling: An Integration of the Best Features of Exploratory and Confirmatory Factor Analysis." *Annual Review of Clinical Psychology* 10: 85–110.
Maulana, R., M. Helms-Lorenz, and W. van de Grift. 2015. "Development and Evaluation of a Questionnaire Measuring Pre-service Teachers' Teaching Behaviour: A Rasch Modelling Approach." *School Effectiveness and School Improvement* 26: 169–194.
McCarthy, C. J., R. G. Lambert, S. Lineback, P. Fitchett, and P. G. Baddouh. 2016. "Assessing Teacher Appraisals and Stress in the Classroom: Review of the Classroom Appraisal of Resources and Demands." *Educational Psychology Review* 28: 577–603.
Mearns, J., and J. E. Cain. 2003. "Relationships between Teachers' Occupational Stress and Their Burnout and Distress: Roles of Coping and Negative Mood Regulation Expectancies." *Anxiety, Stress & Coping* 16 (1): 71–82.
Ministerie van Onderwijs, Cultuur en Wetenschap [Dutch Ministry of Education]. 2014. "Kerncijfers 2009–2013, Onderwijs, Cultuur en Wetenschap [Key Figures 2009–2013, Education, Culture and Science]." https://www.rijksoverheid.nl/documenten/jaarverslagen/2014/05/21/ocw-kerncijfers.
Mokken, R. J. 1971. *A Theory and Procedure of Scale Analysis*. The Hague: Mouton.
Molenaar, I. W., and K. Sijtsma. 2000. *User's Manual MSP5 for Windows*. Groningen: iecProGAMMA.
Notelaers, G., H. De Witte, M. J. P. M. Van Veldhoven, and J. K. Vermunt. 2007. "Construction and Validation of the Short Inventory to Monitor Psychosocial Hazards." *Médecine du Travail et Ergonomie* 44 (1/4): 11–17.
Payne, M. A., and A. Furnham. 1987. "Dimensions of Occupational Stress in West Indian Secondary School Teachers." *British Journal of Educational Psychology* 57 (2): 141–150.
Ronfeldt, M., S. Loeb, and J. Wyckoff. 2013. "How Teacher Turnover Harms Student Achievement." *American Educational Research Journal* 50 (1): 4–36.
Roorda, D. L., H. M. Koomen, J. L. Spilt, and F. J. Oort. 2011. "The Influence of Affective Teacher–student Relationships on Students' School Engagement and Achievement a Meta-analytic Approach." *Review of Educational Research* 81 (4): 493–529.

Rudow, B. 1999. *Stress and Burnout in the Teaching Profession: European Studies, Issues, and Research Perspectives*. New York, NY: Cambridge University Press.

Shen, J. 1997. "Teacher Retention and Dropout in Public Schools: Evidence from SASS91." *The Journal of Educational Research* 91 (2): 81–88.

Sparks, K., and C. L. Cooper. 1999. "Occupational Differences in the Work-strain Relationship: Towards the Use of Situation-specific Models." *Journal of Occupational and Organizational Psychology* 72 (2): 219–229.

Struyven, K., and G. Vanthournout. 2014. "Teachers' Exit Decisions: An Investigation into the Reasons why Newly Qualified Teachers Fail to Enter the Teaching Profession or Why Those Who Do Enter Do Not Continue Teaching." *Teaching and Teacher Education* 43: 37–45.

Van de Grift, W. 2007. "Quality of Teaching in Four European Countries: A Review of the Literature and Application of an Assessment Instrument." *Educational Research* 49 (2): 127–152.

Van de Grift, W. J. 2014. "Measuring Teaching Quality in Several European Countries." *School Effectiveness and School Improvement* 25 (3): 295–311.

Van Veldhoven, M. 1996. "Psychosociale arbeidsbelasting en werkstress [Psycho-social workpressure and workstress]." Doctoral thesis, Rijksuniversiteit Groningen, Groningen, Netherlands.

Van Veldhoven, M., and T. F. Meijman. 1994. *Het meten van psychosociale arbeidsbelasting met een vragenlijst: de vragenlijst beleving en beoordeling van de arbeid (VBBA)* [Measuring Psycho-social Workpressure and Workstress with a Questionnaire: The Questionnaire on the Experience and Evaluation of Work (QEEW)]. Amsterdam: Nederlands Instituut voor Arbeidsomstandigheden.

Van Veldhoven, M., J. Prins, P. van der Laken, and L. Dijkstra. 2015. *QEEW2. 0; 42 Short Scales for Survey Research on Work, Well-being and Performance*. Amsterdam: SKB.

Wellborn, J., J. Connell, E. Skinner, and L. Pierson. 1992. *Teacher as Social Context (TASC)*. Rochester, NY: University of Rochester.

Wilhelm, K., J. Dewhurst-Savellis, and G. Parker. 2000. "Teacher Stress? An Analysis of Why Teachers Leave and Why They Stay." *Teachers and Teaching: Theory and Practice* 6 (3): 291–304.

Appendices

Appendix A. Overview stress causes and stress responses scales used in each sample, and number of items per scale which are used in each sample and the final version of the QEEW-BT

		Scale	Sample 1 QEEW (N items)	Sample 2 QEEW-BT (N items)	Final version QEEW-BT (N items)	H (*t*) QEEW-BT	Rho (*t*) QEEW-BT	Cron-bach's alpha QEEW-BT
Stress causes	Step 1	Pace and amount of work	11	6	6	.58	.83	.85
		Emotional workload	7	6	6	.45	.74	.70
		Mental workload	7	4	4	.57	.76	.75
		Role conflict	6	6	6	.49	.76	.73
		Poor relationship with colleagues	9	6	6	.59	.85	.85
		Poor relationship with supervisor	9	5	5	.71	.89	.88
		Lack of role clarity	5	4	4	.61	.81	.82
		Lack of feedback	7	4	4	.65	.83	.82
		Lack of job variety	6	4	4	.68	.86	.81
		Lack of learning opportunities	4	4	4	.59	.83	.83
		Lack of job autonomy	11	6	6	.55	.83	.80
		Lack of participation	8	4	4	.54	.74	.75
		Lack of communication	4	4	4	.67	.83	.83
	Step 2	Uncertainty about the future	4	4	4	.79	.93	.93
		Physical effort	7					
		Lack of possibilities for contact	4					
		Lack of career opportunities	4					
		Lack of remuneration	5					
		Changes in tasks	5					
	Step 4	Student misbehaviour*		13	12	.52	.90	.89
		Poor relationship with students 1*		8	8	.46	.85	.82
		Poor relationship with students 2*		5	5	.44	.76	.76
		Poor school climate*		11	7	.40	.78	.77
		Education specific workload*		9	9	.36	.80	.80
Stress responses	Step 1	Lack of work pleasure	9	8	8	.42	.74	.74
		Lack of organizational commitment	8	8	8	.45	.74	.75
		Turnover	4	4	4	.82	.86	.75
		Need for recovery	11	8	8	.45	.82	.79
		Rumination	4	4	4	.76	.83	.77
		Emotional reactions during work	12	9	9	.45	.78	.70
		Poor sleep quality	14	7	7	.57	.80	.77
		Tiredness during work	16	14	14	.60	.94	.94

*Rho (*t*), H (*t*) and Cronbach's alpha for the new scales are based on sample 2. For all the other scales they are based on Sample 1.

Appendix B


Van Veldhoven (1996) QEEW
Structure 1
PCA BTs sample 1 QEEW
Structure 2a
PCA BTs sample 2 QEEW-BT
Structure 3a
Stress causes
High psychological task demands
Pace and amount of work
Emotional workload
Mental workload
Role conflict
Education specific workload
Social and organizational aspects
Poor relationship with colleagues
Poor relationship with supervisor
Role conflict
Lack of role clarity
Lack of feedback
Lack of communication
Negative social and organizational aspects
Poor school climate
Lack of development opportunities
Lack of job variety
Lack of learning opportunities
Lack of job autonomy
Negative organisational aspects
Lack of participation
Lack of influence
Negative student aspects
Poor relationship with students 1
Poor relationship with students 2
Student misbehaviour
Stress responses
Structure 2b
Structure 3b
Discontent
Lack of work pleasure
Lack of organizational commitment
Turnover
Tension
Need for recovery
Rumination
Emotional reactions during work
Poor sleep quality
Tiredness during work
Negative Emotions
.67
.73
.63
.54
.61
.68
.79
.74
.81
.74
.59
.69
.78
.83
.81
.71
.58
.59
.71
.75
.60
.80
.73
.73
.54
.76
.85
.75
.87
.65
.72
.82
.80
.85
.78
.65
.79
.70
.74
.80
.72
.74
.93
.90

Appendix C. Explanation of the technical and statistical terms used in the current study

Term	Explanation
Confirmatory Factor Analysis	Factor analysis is a method of modelling the covariation among a set of observed variables as a function of one or more latent constructs (constructs that cannot directly be measured like stress). This technique can be used to construct a questionnaire to measure an underlying construct. Two broad classes of factor analytic methods are Confirmatory Factor Analysis (CFA) and Exploratory Factor Analysis (EFA). Whereas both techniques model the observed covariation among variables as a function of latent construct, in EFA the purpose of the models is to identify the latent constructs or to generate hypotheses about their possible structures, whereas the purpose of CFA is to evaluate hypothesized structures of the latent constructs and/or to develop a better understanding of such structures. Therefore, CFA should only be used if the structure of the variables has been previously studied using Exploratory Factor Analysis (EFA) and independent source of data (Bandalos and Finney 2010). In this study CFA was used to examine the underlying latent constructs of the stress variables for the beginning teacher context. CFA was appropriate to use as the structure of the questionnaire was previously studied using EFA and independent source of data by authors (Van Veldhoven 1996; Van Veldhoven and Meijman 1994).
Principal Component Analysis	Closely related to EFA is Principal Component Analysis (PCA). This is a method for reducing the dimensionality of a set of observed variables through the creation of an optimum number of weighted composites. This technique is mostly used as a tool in exploratory data analysis and can be used to construct a questionnaire or to explore the new factor structure of a questionnaire when new items have been added (Field 2013). In this study PCAs were conducted to explore the factor structure of the QEEW and later the QEEW-BT.
Mokken Scaling *Unidimensionality* *Local independency* *Non-intersecting item response curves* *H(i) score* *H(t) score* *Rho (t)*	Mokken scaling is a psychometric method which can be used for data reduction. For a scale to be a Mokken scale it has to meet several assumptions. Firstly, the assumption of *unidimensionality*. A scale is unidimensional when the items of the scale measure the same latent trait. Secondly, the assumption of *local independency*. This assumption is met when the response to one item does not influence the response to another item, except for an influence that can be explained from the latent variable which is being measured with the set of items. Further, for a Mokken scale to meet the assumption of the double monoticity model, the last assumption states that *the response curves of the items are not allowed to intersect*. The extent to which a set of items is unidimensional is given by the Loevinger's coefficient (H). H can be calculated for individual items *H(i)* and for the overall set of items *H(t)*. A H(i) value of minimal 0.3 and a H(t) value of minimal 0.35 is acceptable (Mokken 1971). The extent to which a set of items is reliable is given by the *Rho(t)*. Rho (t) is comparable to the way internal consistency is calculated with Cronbach's alpha.
Cronbach's alpha	Cronbach's alpha, α, is the most common measure of scale reliability. It is a measure of internal consistency, meaning, a how closely related a set of items are as a group. Generally a value of above 0.80 is great and a value below .70 is unacceptable (Field 2013). However, Cortina (1993) argued that such general guidelines need to be used with caution as the value of α depends on the number of items on the scale. As the number of items on the scale increases, the α will increase.
Spearman's rho	Spearman's rho is a non-parametric statistic which can be used to calculate the correlation between variables when the data has violated parametric assumptions such as non-normally distributed data. Spearman's test works by first ranking the data, and then applying Pearson's equation to those ranks (Field 2013). In this study the data violated the assumption of normally distributed data and therefore Spearman's rho was used for the calculation of the correlations between the variables.