



# Does the School Context *Really* Matter for Teacher Burnout? Review of Existing Multilevel Teacher Burnout Research and Results From the Teaching and Learning International Survey 2018 in the Flemish- and French-Speaking Communities of Belgium

Filip Van Droogenbroeck<sup>1</sup> , Bram Spruyt<sup>1</sup> , Valérie Quittre<sup>2</sup>, and Dominique Lafontaine<sup>2</sup>

It is widely believed that the school context plays a crucial role in teacher burnout. Against that background, we (1) critically review existing empirical multilevel studies on teacher burnout and (2) use data from the Teaching and Learning International Survey (TALIS) 2018 to assess the school-level variance and its correlates in emotional exhaustion, cynical depersonalization, and personal accomplishment in 2,300 primary (183 schools) and 2,700 lower secondary (190 schools) teachers in the Flemish-speaking community of Belgium, and 2,135 lower secondary (120 schools) teachers in the French-speaking community of Belgium. Our results reveal that (1) conceptual confusion exists surrounding school-context variables and (2) the between-school variance in teacher burnout is small. Implications for educational policy and teacher burnout research are discussed.

**Keywords:** emotional exhaustion; hierarchical linear modeling; school context; survey research; teacher burnout

Teacher burnout is considered to be the final stage in a chain reaction caused by chronic occupational stress. Teachers suffering from burnout experience decreased well-being, motivation, and performance, all of which influence their relationships with their students and colleagues (Maslach & Leiter, 1999). Teacher burnout is a problem that has serious consequences, both for teachers' careers (higher risk of absenteeism, lower commitment, lower job satisfaction, and ultimately leaving the profession entirely) and their students (lower interaction, poorer teaching, and lower academic achievement; Vandenberghe & Huberman, 1999).

Twenty years ago, Maslach and Leiter (1999) already drew attention to the role of the school context (e.g., urban vs. rural, student composition, principal leadership, organization of work, etc.) in teacher burnout. In their influential review study, Maslach et al. (2001) concluded that "situational and organizational factors play a bigger role in burnout than individual ones" (p. 418).

Although it seems self-evident that the school context plays an important role in teacher burnout, it is surprising that in 40 years of teacher burnout research, only a few studies exist that use an appropriate multilevel approach that allows for the simultaneous investigation of the individual- and school-level variance (Subramanian et al., 2009). Nevertheless, a correct assessment of the relative importance of the individual and school context is crucial in determining teacher burnout prevention strategies. Substantial between-school differences in teacher burnout would strongly indicate that school organizational factors could be related to teacher burnout and prevention strategies should be directed at this level. The absence of such differences, however, would indicate that teachers' burnout problems might be less related to the specific features of the schools' organization but rather originate in individual teacher's coping styles.

<sup>1</sup>Vrije Universiteit Brussel, Brussel

<sup>2</sup>Université de Liège, Liège

Against that background, our aim is twofold. First, we review the existing multilevel studies on teacher burnout that investigate school differences, both with regard to their conclusions and their methodology. Our central research question in the first part of the study is “Does existing teacher burnout research show substantial between-school variation in teacher burnout?” This review leads to the conclusion that multilevel teacher burnout research is scarce, it often suffers from methodological limitations, and the importance of the school context for teacher burnout is rather limited. Second, we investigate the school-level variance and its correlates in emotional exhaustion, cynical depersonalization, and personal accomplishment using data gathered in the Teaching and Learning International Survey (TALIS) 2018. This provides us with data from 2,300 primary (183 schools) and 2,700 lower secondary (190 schools) teachers in the Flemish-speaking community of Belgium, and 2,135 lower secondary (120 schools) teachers in the French-speaking community of Belgium.

We contribute to the literature in two ways. First, we provide an overview of the existing multilevel teacher burnout literature that is exhaustive to the best of our knowledge. Throughout this process, we aim to clarify some of the conceptual confusion regarding school-context predictors in teacher burnout research. Second, we contribute large-scale high-quality empirical data on teacher burnout. As this data was gathered in three different educational contexts among representative samples of teachers and schools, it allows us to assess the relative importance of the school context on teacher burnout.

## Teacher Burnout

### *What Is Teacher Burnout?*

Burnout is commonly conceptualized as a multidimensional construct, consisting of three linked components: emotional exhaustion, depersonalization, and reduced personal accomplishment (Maslach & Jackson, 1986). *Emotional exhaustion* refers to the feelings of fatigue that arise when individual teachers feel overextended and their emotional energy becomes drained. To cope with this exhaustion, people distance themselves from work psychologically. This results in *depersonalization*, which is reflected by indifferent, impersonal, and dehumanized attitudes toward other people, usually toward the recipients of one’s care (e.g., pupils). Because such attitudes cause work performance and interpersonal relationships to deteriorate, they often lead to a sense of *reduced personal accomplishment*, which arises from negative self-evaluation regarding one’s work-related achievements and is often accompanied by feelings of inadequacy and poor professional self-esteem.

Forty years of research has revealed many possible predictors of teacher burnout. Although several typologies exist to categorize the determinants of burnout, the distinction between individual, organizational, and transactional factors as proposed by Chang (2009) is most useful to our purpose and clarifies some of the conceptual confusion surrounding the work-related antecedents of burnout.

According to Chang (2009), *individual* factors revolve around the question “Who becomes burned out?” and refer to variables

such as age, gender, work experience, and so on. *Organizational* factors revolve around the question “In what context do teachers become burned out?” and include variables such as social support, work demands, student discipline problems, leadership style, participation in school decision making, working conditions, classroom size, student composition, and so on. Note that these are structural aspects of the school and thus should be measured directly at the school level or as emergent collective properties of the perceptions of teachers within a school. From a policy perspective, organizational factors can be more readily directly influenced by policymakers compared with individual or transactional factors. *Transactional* factors are measured at the teacher level and are distinguished from individual and organizational factors. They reflect *how teachers experience and perceive aspects of their work environment* such as social support, student misbehavior, teacher self-efficacy, workload, and so on. To be clear, in our conceptualization, transactional factors become organizational factors once they are aggregated at the school level. Teacher burnout studies often confuse transactional factors with organizational factors. For example, in some cases, teachers’ individual perception of collective school culture or school leadership style is conceptualized as a school-context variable whereas it was measured and analyzed only at the individual level. As we illustrate in the section “Is Teacher Burnout Related to the School Context?” such practices may lead to conceptual and statistical problems.

### *Why Would Teachers’ Burnout Depend on the School Context?*

Burnout is mainly considered to be a job-related phenomenon caused by prolonged stress in the work environment (Schaufeli & Enzmann, 1998). Therefore, there are good reasons to expect that the environment in which teachers work plays an important role in teacher burnout.

Teachers from the same school share more common tasks and goals with each other than they do with teachers from different schools, and are exposed to the same workplace environment (Organisation for Economic Co-operation and Development [OECD], 2019). An organizational perspective on teacher burnout starts from the conviction that the more schools constitute supportive, collegial, and healthy work environments, the lower the risk on burnout among individual teachers.

When this is applied to the job demands and resources (JD-R) model, arguably one of the most popular models to explain teacher burnout, the following rationale arises. In the JD-R model (Bakker & Demerouti, 2007), all work characteristics are classified into two global categories: job demands and job resources. The model states that excessive job demands and lack of job resources cause professional burnout to develop. Clearly, job demands and resources can also be understood at the transactional and organizational level. Through the lens of the JD-R model, between-school differences in teacher burnout can be expected because teachers in the same school share similar organizational demands and resource scarcity. Though other theories could be used, such as social disorganization theory (Sampson & Groves, 1989) or conservation of resources theory (Hobfoll, 1989), they all share the same core argument, namely, that the

degree to which the organizational context is unfavorable can increase individual teacher's stress and ultimately lead to burnout. After having clarified the potential importance of the school level, we now turn to our core research question: "How much evidence do we have that the school level matters in explaining teacher burnout?" To answer this question, we first performed an exhaustive literature review.

## Overview of Multilevel Studies on Teacher Burnout

### *Research Strategy*

As a first step in gathering existing research, we performed a comprehensive search on Scopus using terms such as "teach\* burn\*," "teach\* emotional exhaustion," "teach\* depersonalization," and "teach\* personal accomplishment." This generated a set of 4,143 articles published between 1974 and 2020. Next, articles were filtered based on the presence of keywords such as "\*multi\*," "\*hierarchical linear model\*," "\*school\*level\*," and "\*context\*" in the title, abstract, and keywords, which resulted in a list of 583 articles. These were manually checked. A comprehensive database search was additionally performed (e.g., Web of Science, Elsevier ScienceDirect, ERIC, JSTOR, Google Scholar, etc.) using a snowball approach. Here, we used keywords such as "teacher burnout" and "teacher burnout multilevel." Studies were included when they met the following criteria: (1) studies had to be in English and peer-reviewed, (2) a burnout measure had to be included as a dependent variable, and (3) a multilevel analysis approach among teachers had to be used. As we will discuss below, the latter is a crucial prerequisite to correctly analyze individual and school-level determinants of teacher burnout simultaneously. We subsequently searched for additional relevant studies using the reference list. Our search resulted in a final total of 14 teacher burnout studies multilevel (see Table 1). The studies are summarized in the supplemental appendix (available on the journal website).

### *Is Teacher Burnout Related to the School Context?*

A review of the studies leads to four conclusions. First, the number of studies that use a multilevel design to study teacher burnout and how it relates to the school context is extremely low considering the 40-year history of this field and the large number of teacher burnout studies that have been carried out (after further data cleaning, our conservative estimate is 1,645 studies).

Many studies investigate the "school context" but only use transactional variables (e.g., Ainsworth & Oldfield, 2019; Cano-García et al., 2005; Conley & You, 2018; Parrello et al., 2019; Pietarinen et al., 2013; Skaalvik & Skaalvik, 2009, 2017). These studies rely on individual-level teacher data and apply single-level analysis to investigate the school context. As these studies made important contributions to the literature, we do not single them out as examples of bad research but refer to them only to illustrate the conceptual confusion surrounding school-context variables in teacher burnout research.

In general, using a single-level approach to investigate the school context can be problematic for two reasons. The first problem is conceptual. Education systems are characterized by a multileveled nature, where teachers are embedded within schools. Reducing this complexity to a single-level investigation is statistically prone to "ecological" and "individualistic" fallacies, where data are analyzed at one level but conclusions are formulated at another (Hox et al., 2017). In an ecological fallacy, invalid inferences about individuals are made based on aggregated group-level data. The opposite and lesser-known individualistic fallacy was identified by Alker (1969) as occurring in cases where social scientists try to generalize from individual behavior to collective relationships. Indeed, assuming that the relationships between variables at the teacher level are the same at the school level and vice versa might be misleading.

It is clear, for example, that a research question such as "Does social support reduce teacher burnout at the individual level?" fundamentally differs from "Do teachers in schools characterized by more social support report less teacher burnout?" Problems especially arise when single-level results are used to infer that the school context plays a crucial role in teacher burnout and that schools should change policies according to the findings (e.g., increase social support in schools to prevent teacher burnout, change leadership styles, create supportive environments, etc.). This is a common practice that simply cannot be supported by individual-level data.

The second problem is statistical. Some teacher burnout studies include school-level predictors (e.g., school size, primary vs. secondary, urban vs. rural) in a single-level model (e.g., Skaalvik & Skaalvik, 2009; Steinhardt et al., 2011). This entails a high risk of Type I errors, as the standard errors of the school-level predictors may be severely underestimated, leading to spuriously narrow confidence intervals for the school-level regression coefficients (Merlo et al., 2018; Subramanian et al., 2009). Multilevel models have been specifically designed to avoid this type of problem.

A second conclusion from our literature review is that the teacher burnout studies that did apply multilevel analysis have studied a very wide range of school-level variables (Table 1). This wide range of school-level variables being tested in different studies is not necessarily problematic, but it does imply that the relationship between many school-level variables and teacher burnout has been investigated only once. Therefore, more studies should be welcomed.

Third, the methodological quality of these studies varies strongly. The *first* methodological issue concerns the different criteria that are used to determine when to include school-level variables and whether a multilevel analysis is deemed appropriate. Twelve out of 14 studies reported the intraclass correlation coefficient (ICC) to assess between-school variance. Five studies (Lim & Eo, 2014; McCarthy et al., 2009; McCormick & Barnett, 2011; Ullrich et al., 2012; Van Maele & Van Houtte, 2015) chose not to include school-level variables when the ICC was considered too low or if the school-level variance component was not significant. Clearly, the level at which some studies (Lim & Eo, 2014; McCarthy et al., 2009; McCormick & Barnett, 2011) deem the ICC too low to include school-level

**Table 1**  
**Overview of Teacher Burnout Studies That Used or Tested a Multilevel Approach**

Authors	Sample (Data Year)	Burnout Measure	ICC	School-Context Effects
1. Klusmann et al. (2008)	Germany (2003)	MBI		Standardized coefficients
	1939 mathematics and science teachers	EE	0.010	<i>Principal perspective</i>
	198 secondary education schools			Students discipline (–.02) Teacher morale (.01) <i>Teachers' perspective</i> Principal support (–.04) Cooperation between colleagues (.03) Students discipline (–.11**) <i>Students' perspective</i> Students' SES (–.01) Students' basic cognitive abilities (–.01)
2. McCarthy et al. (2009)	USA: Urban region south east (Unk.)	MBI-ES		
	451 teachers	EE	0.021	Not included
	13 primary education schools	DP	0.012	Not included
		PA	0.005	Not included
3. McCormick and Barnett (2011)	Australia: New South Wales (Unk.)	MBI-ES		
	416 teachers	EE	0.028	Not included
	38 secondary education schools	DP	0.068	Not included
		PA	0.000	Not included
4. Pas et al. (2012)	USA: Maryland (2007–2009)	MBI		Standardized coefficients
	600 teachers	EE	0.048	Average overall organizational health (0.15)
	31 primary education schools		(personal communication)	Student mobility rate (0.00) Student suspension rate (0.01) Student enrollment (0.05) Principal turnover (–0.07)
5. Ross et al. (2012)	USA: Oregon (2009)	MBI-ES		Unstandardized coefficients
	184 teachers	EE	Unk.	Percentage of students receiving free and reduced-price lunch (SES) (4.62**)
	40 primary education schools	DP	Unk.	Implementation level of school-wide positive behavioral interventions and supports school (SET) (–8.57***)
		PA	Unk.	SES (11.39**) SET (2.00) SES × SET (–9.68*) SES (–22.93***) SET (–5.01 <sup>†</sup> ) SES × SET (23.02***)
6. Ullrich et al. (2012)	Germany: Baden-Württemberg (Unk.)	MBI	Range 0.000016–0.00003	
	460 teachers	EE	Unk.	Not included
	62 primary education schools (49 general education primary schools, 13 special education elementary schools)	DP	Unk.	Not included
		PA	Unk.	Not included
7. González-Morales et al. (2012)	Spain (Unk.)	MBI-GS		Unstandardized coefficients
	555 teachers	EE	T2 = 0.062	Primary/secondary school (0.08) Teacher–students ratio (–6.0) Absenteeism rate (–0.02) Quality of school facilities (–0.03) Perceived collective EE T1 (.33*)

*(continued)*

Table 1 (continued)

Authors	Sample (Data Year)	Burnout Measure	ICC	School-Context Effects
	100 schools (63 primary, 37 secondary education schools)	DP	T2 = 0.082	Primary/secondary school (0.20) Teacher–students ratio (–6.6*) Absenteeism rate (–0.01) Quality of school facilities (–0.04) Perceived collective EE T1 (0.36)
8. Lim and Eo (2014)	South Korea: Kangwon province (Unk.) 367 teachers	MBI-ES EE	Six measures including EE-DP-PA range 0.020–0.080	Not included
	24 public middle schools	DP PA		Not included Not included
9. Van Maele and Van Houtte (2015)	Belgium: Flemish-speaking community (2008–2009) 673 teachers 58 primary education schools	MBI-ES EE DP PA	0.015 0.050 0.000 0.020	Not included Not included Not included
10. Kim et al. (2017)	USA: Michigan and Indiana (2007–2009) 171 early career teachers 84 primary and middle education schools	Nine item unspecified burnout scale	T1 = 0.193 T2 = 0.230	Unstandardized coefficients Elementary school (–0.063**) Students eligible for free lunch (0.046**) Organizational exposure (mean burnout) (0.90***)
11. O'Brennan et al. (2017)	USA: Maryland (2012) 3225 teachers and paraprofessionals (75% were teachers) 58 secondary education schools (grades 9–12)	MBI EE	0.056	Standardized coefficients Student–teacher ratio (0.006) Suspension ratio (0.004*) Free and reduced-price meals rate (0.000) School physical orderliness (0.031) Urbanicity (0.012) Positive behavioral interventions and supports school (0.006)
12. Zheng et al. (2017)	China (2013) 8,563 teachers 583 secondary education schools (Grade 8)	Nine-item burnout scale inspired on MBI	0.150	Unstandardized coefficients <i>Principal perspective</i> School size (–0.00) School location (ref. rural) City (0.08*) County (0.03) General teaching resources (–0.01) Student–teacher ratio (0.01*) Visibility and direct participation (0.02) Instruction organization (–0.00) Internal environment (–0.00) Planning and personnel (0.00) External relations (0.03*) <i>Teachers' perspective</i> School size (–0.00) School location (ref. rural) City (0.10*) County (0.07) General teaching resources (0.02) Student–teacher ratio (0.01*) Visibility and direct participation (–0.16***) Instruction and curriculum (–0.07**) Organization and management (–0.06*)

(continued)

Table 1 (continued)

Authors	Sample (Data Year)	Burnout Measure	ICC	School-Context Effects
13. Ford et al. (2019)	USA: urban district Midwestern state (2017)	MBI	0.070	Standardized coefficients
	781 teachers 73 primary and secondary education schools	Composite teacher burnout measure		Collective teacher efficacy (−0.37***) Organizational support for teachers' psychological needs (−0.255***)
14. Shackleton et al. (2019)	United Kingdom: South-East England (2014)	MBI	0.050	Unstandardized coefficients
	2,278 staff members (teachers, teaching assistants, heads of year, heads of department, senior managers, other) 39 secondary education schools	EE		School type (ref. Academy-converter) Voluntary (−1.22) Community school (0.19) Academy-sponsor-led (3.69) Foundation school (0.61) School quality inspection report Good (0.30) Requires improvement (1.99) Single-sex school status (ref. mixed) All girls (0.56) All boys (1.50) Size of school (−0.10) Income Domain Affecting Children Index (IDACI) score (0.06*) Teacher perceived safety (ref. All the time) Most of the time (6.23**) Some of the time/never (11.74**) Teacher perceived support (ref. very well) Quite well (4.55**) Not very well (9.70**) Not at all (13.92**) Student–teacher ratio (−0.21) Student attitude to learning (0.98**) Free school meals (0.08**) Proportion of Special Education Needs (SEN) Students (0.30**) Proportion of English as an Additional Language (EAL) (0.04 <sup>†</sup> )
		DP	0.050	School type (ref. Academy-converter) Voluntary (0.25) Community school (0.16) Academy-sponsor-led (1.97**) Foundation school (0.92*) School quality inspection report Ofsted rating Good (0.52) Requires improvement (1.75*) Single-sex school status (ref. mixed) All girls (−0.87 <sup>†</sup> ) All boys (1.06) Size of school (0.00) IDACI score (0.02*) Teacher perceived safety (ref. All the time) Most of the time (2.31**) Some of the time/never (4.80**) Teacher perceived support (ref. very well) Quite well (1.68**) Not very well (3.87**) Not at all (5.50**) Student–teacher ratio (−0.14) Student attitude to learning (0.48**) Free school meals (0.03*) Proportion of SEN students (0.15**) EAL (0.01)

(continued)

Table 1 (continued)

Authors	Sample (Data Year)	Burnout Measure	ICC	School-Context Effects
		PA	0.020	School type (ref. Academy-converter) Voluntary (−2.61**) Community school (−0.71) Academy sponsor-led (−1.72**) Foundation school (−0.74) School quality inspection report Ofsted rating Good (−0.69) Requires improvement (−0.77) Single-sex school status (ref. mixed) All girls (0.44) All boys (−0.81) Size of school (0.03) IDACI score (−0.02*) Teacher perceived safety (ref. All the time) Most of the time (−1.80**) Some of the time/never (−4.31**) Teacher perceived support (ref. Very well) Quite well (−2.50**) Not very well (−4.02**) Not at all (−3.62**) Student-teacher ratio (0.31*) Student attitude to learning (−0.57**) Free school meals (−0.02 <sup>†</sup> ) Proportion of SEN students (−0.06) EAL (0.00)

*Note.* ICC = intraclass correlation coefficient; MBI = Maslach Burnout Inventory; MBI-GS = MBI-General Survey; MBI-ES = MBI-Educator Survey; EE = emotional exhaustion; DP = depersonalization; PA = personal accomplishment; Unk. = unknown, USA = United States of America; SES = socioeconomic status; SET = school-wide evaluation tool.

<sup>†</sup> $p < .10$ . \* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

variables or to use multilevel analysis is considered acceptable in other studies (Ford et al., 2019; González-Morales et al., 2011; Klusmann et al., 2008; O'Brennan et al., 2017; Pas et al., 2012; Shackleton et al., 2019). There is no consensus in the literature as to what level of ICC is suitable to warrant using multilevel analysis or to include school-level variables (Hox et al., 2017). However, research has shown that not taking the hierarchical structure of the data into account can lead to Type-I errors, even when the ICCs are very small (Musca et al., 2011). Scholars recommend to investigate the design effect rather than the ICC to justify not taking the clustered structure of the data into account (Lai & Kwok, 2015). The design effect is calculated as

$$\text{Design effect} = 1 + (n - 1) \times \text{ICC}$$

where  $n$  is the average group size. The design effect reflects how much the standard errors are underestimated when a hierarchical sample is treated as a simple random sample (Maas & Hox, 2005). As a rule of thumb, it is suggested that if the design effect is less than two, a single-level analysis of multilevel data does not seem to lead to misleading results but *only if* (1) the analysis is solely interested in individual-level relationships, (2) teacher-level predictors do not have any school-level effects on the outcome, and (3) the effects of the predictors do not vary across schools. When researchers are interested in school-level

predictors, they should *always* use techniques that adequately take into account the complex data structure, unless there are not enough schools (<20) or the design effect is smaller than 1.1 (Lai & Kwok, 2015; Maas & Hox, 2005).

Closely related to the previous argument, the *second* methodological issue is that when a significant school-level effect is found, *the magnitude of that effect* is not always taken into account. This is a great limitation because a paradoxical relationship exists between ICCs and school-level predictors, where smaller ICCs actually lead to more precise estimates of the latter. For this reason, Merlo et al. (2018) stress that it is crucial to simultaneously look at the magnitude of the general contextual effect (e.g., ICC) and specific contextual effects (e.g., standardized coefficient of school-level predictors) to assess the relative contribution of the school level. Few scholars are aware of this. Applied to teacher burnout, this would mean that when the school context is less relevant (e.g., a small ICC in the unconditional model before including school-context predictors), it will be easier to find statistically significant effects of school-level predictors because of the increased effective sample size. Thus, even if researchers use a multilevel approach but solely focus on specific school-level predictors without taking the magnitude of the ICC into account, they run the risk of reporting trivially small but statistically significant effects. Accordingly, they might conclude that the school level is relevant when in fact it is not.

Across the selected studies, the proportion of the variance situated at the school level ranged between 1.0% and 6.2% for emotional exhaustion; 0.0% and 8.2% for depersonalization; and 0.0% and 3.0% for personal accomplishment. One study (Kim et al., 2017) used an unspecified nine-item burnout scale and reported an ICC of 23.0% but based its results on 171 early career teachers in 84 primary and middle schools, which resulted in an average of two teachers per school. Such a low number of teachers per school will obviously inflate the between-school variance. Two studies (Ford et al., 2019; Zheng et al., 2017) used a composite teacher burnout measure based on the Maslach Burnout Inventory (MBI) with an ICC of 15.0% and 7.0%, respectively. Teacher burnout in the MBI is conceptualized as a multidimensional construct, and previous research indicated that variables are differentially related to the three burnout dimensions. For this reason, it is advised that the three dimensions of burnout should not be combined in a single measure (Byrne, 1994; Skaalvik & Skaalvik, 2009; Van Droogenbroeck et al., 2014).

Regarding the school-level predictors, 10 out of 14 studies included school-level variables. Four of these studies (Ford et al., 2019; Klusmann et al., 2008; O'Brennan et al., 2017; Pas et al., 2012) reported standardized coefficients for the school-level variables. Eight studies (Ford et al., 2019; González-Morales et al., 2011; Kim et al., 2017; Klusmann et al., 2008; O'Brennan et al., 2017; Ross et al., 2012; Shackleton et al., 2019; Zheng et al., 2017) found significant relationships between school-level variables and a burnout measure, but only three of them (Klusmann et al., 2008; O'Brennan et al., 2017; Zheng et al., 2017) included a measure for the effect size of the school-level variable. The relationship between the coefficients of school-level variables and teacher burnout is discussed in the summary of studies in the supplemental appendix (available on the journal website).

The *third* methodological limitation concerns six studies (Ford et al., 2019; González-Morales et al., 2011; Kim et al., 2017; O'Brennan et al., 2017; Shackleton et al., 2019; Ullrich et al., 2012) where different educational levels (e.g., primary and secondary education), type of schools (e.g., general and special education schools), or types of staff (e.g., teachers, paraprofessionals, senior managers, heads of departments, other) are lumped together. This is problematic because it is likely to artificially increase the variance in teacher burnout, as differences between educational levels are likely to differ more than differences between schools. *Fourth*, problems of generalizability arise at the school and population level when only a small number of teachers from a small number of schools are included in the analysis. The number of schools in the reviewed studies ranged from 13 to 583, whereas the average number of teachers in a school ranged from two to 58 teachers. None of the studies used a two-stage probability sampling design where teachers were randomly selected from randomly selected schools from a school population register.

### *New Study*

The core finding from the review above is that (1) in general between-school differences in teacher burnout are small but (2)

it also signals the need for more research that takes the strengths and weaknesses of the previous studies into account. Therefore, in the second part of this article, we use large-scale high-quality data on teacher burnout to assess how much of the total variance in individual-level teacher burnout can be attributed to the school level in three educational contexts.

At the school level, our review reveals a large number of variables that are potentially relevant to teacher burnout. Instead of including dozens of school-level variables to explain a potentially low amount of between-school variance, we will test three school-level job demands and resources that each covers a different and important facet of the organizational context. Following the job demands-resources model, we predict that burnout will be higher for teachers working in more challenging and demanding schools. We examined two commonly used school-level job demands: (1) the proportion of students from socioeconomically disadvantaged homes, which is a widely used indicator that reflects not only the school composition but also the reality that schools generally do not choose their pupils and (2) the frequency of intimidation or verbal abuse of teachers or staff occurring in the school, which is a proxy for the social aspects (interactions with students) of the school (Klusmann et al., 2008; Ross et al., 2012; Shackleton et al., 2019). In terms of resources, schools that form supportive and collegial environments are expected to serve as a buffer against teacher burnout (Ford et al., 2019; Shackleton et al., 2019; Zheng et al., 2017). We measure such a school-level job resource by the presence of a collaborative school climate, which measures whether teachers experience a shared responsibility for school issues and are involved in decision making (Zheng et al., 2017).

At the individual level, control variables are included that are known to be related to teacher burnout. Based on existing meta-analyses of teacher burnout (e.g., Alarcon, 2011; Lee & Ashforth, 1996), we selected demographics (gender and work experience), job resources (perceived teacher support, perceived collaborative school climate, self-efficacy), and job demands (perceived workload and perceived stress from intimidation by students). We expect that the individual- and school-level job demands will be negatively related to emotional exhaustion and depersonalization while positively related to personal accomplishment. On the other hand, individual- and school-level job resources are expected to create a buffer against emotional exhaustion and cynical depersonalization and strengthen personal accomplishment.

## **Method**

### *Data and Sample*

Data were used from TALIS 2018, which was organized by the OECD. A total of 48 countries and economies participated in the third wave of TALIS 2018 (<http://www.oecd.org/education/talis/>). TALIS is the largest cross-national data collection among teachers and principals for a wide range of themes concerning their working life (e.g., job conditions, school climate, teaching practices, etc.).

TALIS uses a stratified two-stage probability sampling design, which is managed by Statistics Canada. Participating countries/regions were first asked to provide a complete list of schools. Schools were selected using systematic random sampling with



probability proportional to the size of teachers within explicit strata. These strata include, for instance, the type of funding and were tailored to the specific context of each country. Teachers were randomly selected in each of the selected schools. TALIS used a hypothetical design effect of 5.2 to derive the expected effective sample size for teachers. It further uses strict procedures to ensure data comparability. Each participating country is required to achieve a response rate of at least 75% of the selected schools and teachers. A school is considered to participate if at least the principal and 50% of the teachers participated in the survey (for a detailed description of the sampling and quality control procedures, see OECD, 2019).

Participating countries and economies were obligated to take part in the core survey that questions teachers and their principals in lower secondary level schools according to the UNESCO International Standard Classification of Education (ISCED) Level 2. In addition to the core survey, one or more international survey options were offered at the primary level (ISCED Level 1) or upper secondary level (ISCED Level 3). All data were collected between March and May 2018. TALIS 2018 participants were able to include national questions. For Belgium, both language communities included measures on teacher burnout. The Belgian language communities have separate governments and are fully responsible for the organization of education and policymaking. This is why the Flemish (Dutch-speaking) and French-speaking community are considered to be separate regions in TALIS and why, for example, only the Flemish community participated in the ISCED1 option. Previous studies have indicated that when compared with other Western countries, socioeconomic and ethnic school segregation in Belgium is high (Jacobs et al., 2009). Educational policies such as parents' unrestricted agency to choose schools and early tracking are thought to play an important role in these dynamics of segregation. For example, middle-class parents tend to avoid schools with a high proportion of immigrant or working-class pupils, even if these schools are in their neighborhood (Agirdag & Van Houtte, 2011).

Response rates for the Flemish-speaking community of Belgium ISCED1 (school-level: 88.6%, teacher-level: 92%) and ISCED2 (school-level: 90.7%, teacher-level: 84.3%) and for Wallonia ISCED2 (school-level: 100%, teacher-level: 89.2%) met the OECD standards. This generated data from 2,672 ISCED1 teachers in 178 schools and 3,198 ISCED2 teachers in 186 schools in the Flemish-speaking community of Belgium, and 2,135 ISCED2 teachers in 120 schools in the French-speaking community of Belgium. The Flemish community participated as an adjudicated subnational entity in the TALIS 2018 survey, which resulted in a higher number of participating schools (OECD, 2019).

## Measures

**Dependent Variable.** Teacher burnout is our dependent variable and was assessed by the Dutch (Horn & Schaufeli, 1998) and French (Vercambre et al., 2009) translations of the Maslach Burnout Inventory–Educators Survey (MBI-ES). Emotional exhaustion (e.g., “I feel used up at the end of the workday”), cynical depersonalization (e.g., “I worry that this job is

hardening me emotionally”), and personal accomplishment (e.g., “I feel I’m positively influencing other people’s lives through my work”) were each measured by four items on a 7-point Likert-type scale ranging from 1 = *never* to 7 = *every day*. These items were used to construct summation scales ranging from 0 to 100 that comprised the three dimensions of teacher burnout. Higher scores indicate more emotional exhaustion, cynical depersonalization, and personal accomplishment. The multidimensional nature of teacher burnout has repeatedly been confirmed, and research has indicated that the three dimensions cannot be combined into a single measure (Byrne, 1994; Langballe, 2006; Schaufeli et al., 1994).

The Cronbach alphas across educational contexts ranged between .87 and .89 for emotional exhaustion, between .70 and .73 for personal accomplishment, and between .59 and .64 for depersonalization and are in line with previous research. It is not uncommon for the reliability of the depersonalization scale to be lower than for the other scales and is known to drop below .70 (Greenglass et al., 1997; Schaufeli et al., 2001; Van Maele & Van Houtte, 2015).

**Independent Variables.** Three school-level indicators were included to assess context effects. Two items from the principal questionnaire were used. School leaders estimated the proportion of students in their schools who come from socioeconomically disadvantaged homes on a 4-point Likert-type scale ranging from 1 = *none* to 4 = *more than 60%*. They further estimated the frequency in which intimidation or verbal abuse of teachers or staff occurred on a 5-point Likert-type scale ranging from 1 = *never* to 5 = *daily*. One variable was created by aggregating teachers' individual responses within schools regarding the presence of a collaborative school culture. To assess the reliability of the school-level construct, the ICC(2) was calculated as follows:  $([\text{mean square between} - \text{mean square within}] / \text{mean square between})$ . Values above 0.60 are considered acceptable to permit aggregation. The ICC(2) across educational contexts ranged between 0.70 and 0.80 for aggregated collaborative school culture. This scale was also included at the individual level, and its operationalization will be discussed in more detail below.

Gender, work experience as a teacher (number of years), self-efficacy, workload, perceived intimidation or verbal abuse by students, perceived teacher support, and collaborative school climate were included as individual-level control variables. TALIS defines self-efficacy as the degree to which teachers believe that they are able to enact certain teaching behaviors that influence students' educational outcomes, such as achievement, interest, and motivation (Ainley & Carstens, 2018). Self-efficacy was measured by 12 items that asked teachers to what extent they could (e.g., “Help students think critically,” “Get students to follow classroom rules”). Teachers' workload was assessed by three items (e.g., “Having too much class preparation”). The extent to which teachers experience intimidation or verbal abuse by students as a source of stress was measured by one item. All the above items were rated on a 4-point Likert-type scale ranging from 1 = *not at all* to 4 = *a lot*.

Perceived teacher support was measured by one item “Teachers can rely on each other.” Perceived collaborative school climate was measured by five items (e.g., “This school has a

**Table 2**  
**Descriptive Statistics for the Dependent Variables: Frequencies, Means, and Standard Deviations**

Variable	Belgium Flemish-Speaking Community						Belgium French-Speaking Community		
	ISCED1			ISCED2			ISCED2		
	M (SD)	Min–Max	N	M (SD)	Min–Max	N	M (SD)	Min–Max	N
Individual-level									
Gender (female)	82.38%		2,662	69.57 %		3,122	68.74 %		2,134
Work experience	16.76 (11.16)	0–44	2,642	15.68 (10.41)	0–44	3,083	14.51 (10.53)	0–43	2,117
Self-efficacy	3.42 (0.39)	1–4	2,517	3.34 (0.40)	1–4	2,945	2.96 (0.43)	1–4	2,023
Workload	2.08 (0.67)	1–4	2,537	2.09 (0.69)	0–4	2,951	2.48 (0.81)	1–4	2,060
Stress by intimidation students	1.54 (0.81)	1–4	2,547	1.66 (0.88)	1–4	2,960	2.03 (1.00)	1–4	2,072
Perceived teacher support	3.42 (0.57)	1–4	2,555	3.27 (0.58)	1–4	2,964	2.92 (0.74)	1–4	2,073
Perceived collaborative school climate	2.96 (0.34)	1–4	2,519	2.79 (0.49)	1–4	2,920	2.68 (0.56)	1–4	2,024
School-level									
Students from socioeconomically disadvantaged homes (PQ)	2.69 (0.85)	1–4	173	2.64 (0.79)	1–4	166	2.80 (0.98)	1–5	115
Intimidation or verbal abuse of teachers or staff (PQ)	1.92 (0.77)	1–5	172	2.15 (0.81)	1–5	168	1.86 (0.74)	1–4	113
Aggregated collaborative school climate	2.96 (0.21)	2.40–3.58	177	2.79 (0.21)	2.15–3.30	182	2.67 (0.27)	1.92–3.38	120

Note. All individual-level variables were group-mean centered except gender. ISCED = International Standard Classification of Education; ISCED1 = primary education; ISCED2 = lower secondary education; PQ = items from principal questionnaire.

**Table 3**  
**Descriptive Statistics for the Dependent Variables: Means, Cronbach's alpha, and Intraclass Correlation Coefficients**

	Belgium Flemish-Speaking Community						Belgium French-Speaking Community		
	ISCED1			ISCED2			ISCED2		
	M (SD)	$\alpha$	ICC	M (SD)	$\alpha$	ICC	M (SD)	$\alpha$	ICC
Emotional exhaustion	42.20 (23.20)	.87	0.054	40.78 (23.61)	.88	0.021	46.87 (26.16)	.89	0.047
Cynical depersonalization	10.21 (13.92)	.59	0.038	14.52 (16.28)	.64	0.040	15.43 (17.67)	.63	0.031
Personal accomplishment	85.69 (11.88)	.70	0.020	78.57 (14.33)	.73	0.006	62.20 (20.24)	.70	0.011
N			2,552			2,949			2,060
Schools			177			182			120

Note. ISCED = International Standard Classification of Education; ISCED1 = primary education; ISCED2 = lower secondary education.

culture of shared responsibility for school issues”). All items included in the last two measures were rated on a 4-point Likert-type scale ranging from 1 = *strongly disagree* to 4 = *strongly agree*. Summation scales were constructed for self-efficacy (range Cronbach's  $\alpha$  = .84-.87), workload (range Cronbach's  $\alpha$  = .63-.78), and collaborative school climate (range Cronbach's  $\alpha$  = .80-.81) ranging from 1 to 4. Confirmatory factor analysis was used to assess how well the empirical data reflected the latent constructs. Goodness-of-fit was evaluated by fit indices such as the root mean square error of approximation and standardized root mean square residual indices where values less than 0.05 indicate a good fit (Browne & Cudeck, 1993). As a rule of

thumb for the comparative fit index and Tucker–Lewis index, values greater than 0.90 are considered acceptable (Hu & Bentler, 1999). All scales in all contexts reached minimum goodness of fit criteria. Detailed goodness-of-fit statistics for each scale is available on request.

### Analysis

Statistical analyses were performed using STATA 16 (Statacorp, 2019). The descriptives are reported in Tables 2 and 3. Multilevel linear regression analysis was used to assess the relationships between school-level variables and emotional exhaustion,

**Table 4**  
**Multilevel Analysis of School-Level Variables (Bivariate) on Emotional Exhaustion, Cynical Depersonalization, and Personal Accomplishment Adjusted for Individual-Level Variables**

	Emotional Exhaustion			Cynical Depersonalization			Personal Accomplishment		
	BFL		BFR	BFL		BFR	BFL		BFR
	ISCED1	ISCED2	ISCED2	ISCED1	ISCED2	ISCED2	ISCED1	ISCED2	ISCED2
Percentage of students from socioeconomically disadvantaged homes (PQ)	0.02	0.06**	0.07*	0.06*	0.12***	0.08**	-0.06*	0.00	-0.02
Intimidation or verbal abuse of teachers or staff (PQ)	0.06*	0.06**	0.07*	0.05*	0.11***	0.08**	-0.03	-0.05*	-0.04
Aggregated collaborative school climate	-0.07**	-0.04*	-0.09**	-0.09***	-0.04†	-0.10***	0.05*	0.05**	-0.01
<i>N</i>	2,390	2,563	1,825	2,387	2,560	1,819	2,387	2,559	1,818
Schools	172	163	113	172	163	113	172	163	113

*Note.* Cell entries are standardized coefficients. School-level variables adjusted for the following individual-level variables: gender, work-experience, self-efficacy, workload, stress by intimidation, perceived teacher support, perceived collaborative school climate. BFL = Belgium Flemish-speaking community; BFR = Belgium French-speaking community; ISCED = International Standard Classification of Education; ISCED1 = primary education; ISCED2 = lower secondary education; PQ = items from principal questionnaire.

† $p < .10$ . \* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

cynical depersonalization, and personal accomplishment. To assess the relative contribution of the school context, we followed Merlo et al.'s (2018) suggestion and simultaneously looked at the magnitude of the general contextual effect (by calculating the ICC) and specific contextual effects (by estimating standardized coefficients). We used standardized coefficients, but other measures of effect sizes are possible for school-level predictors (see Lorah, 2018).

To determine the ICCs (Table 3), we estimated unconditional multilevel models and used the following equation:

$$ICC = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2} \quad (1)$$

where  $\sigma_u^2$  is the between-school variance in individual burnout outcomes, and  $\sigma_e^2$  is the variance between individual burnout outcomes within schools. The ICC, therefore, quantifies the proportion of the variance in teacher burnout situated at the school level. Standardized coefficients for the multilevel models were obtained from the unstandardized coefficients using the formula proposed by Hox et al. (2017, p. 18). Table 4 presents the relationship between the separate school-level variables and teacher burnout, adjusted for the individual-level variables. In Table 5, mutually adjusted models were estimated for school-level variables, adjusted for individual-level variables. Equation (2) summarizes the general full model represented in Table 5.

$$Y_{ij} = \beta_0 + \beta_1 SEX_{ij} + \beta_2 SEFF_{ij} + \beta_3 WRKEXP_{ij} + \beta_4 WRKL_{ij} + \beta_5 INTIMID_{ij} + \beta_6 SUPP_{ij} + \beta_7 COLLSCHCLIM_{ij} + \beta_8 PQ\_SES_j + \beta_9 PQ\_INTIMID_j + \beta_{10} COLLSCHCLIM_j + u_{0j} + \varepsilon_{ij} \quad (2)$$

All individual-level variables (subscript  $ij$ ) were group-mean centered, separating the individual- and second-level variation and allowing for the estimation of purely individual- and group-level effects. This approach maximizes the likelihood of finding significant relationships between school-level variables and the dimensions of burnout (Hox et al., 2017).

## Results

The first objective of our empirical analyses is to determine the school-level variance in burnout in our three samples. The descriptive statistics for all included (in)dependent variables are presented in Tables 2 and 3. The general context effects—or ICCs—for teacher burnout ranged between 2.1% and 5.4% for emotional exhaustion, between 3.1% and 4.0% for depersonalization, and between 0.6% and 2.0% for personal accomplishment. So, similar to the studies in our review, the school level only accounts for a small proportion of the variance in teacher burnout.

Table 4 presents the multilevel models where the school-level variables are analyzed separately, adjusted for teacher characteristics. The multilevel models with mutually adjusted school-level variables, adjusted for teacher characteristics, on the three dimensions of burnout, are presented in Table 5. The models for emotional exhaustion reveal that teachers working in schools with a higher proportion of students from socioeconomically disadvantaged homes report more emotional exhaustion in lower secondary education in the Flemish community (0.06;  $p < .05$ ). This relationship was not significant in primary education in the Flemish community and lower secondary education in the French-speaking community. The frequency of intimidation or verbal abuse of teachers or staff as reported by the principal is positively related to emotional exhaustion in two educational contexts but not in lower secondary education in the French-speaking community of Belgium. Teachers working

**Table 5**  
**Multilevel Analysis of Mutually Adjusted School-Level Variables on Emotional Exhaustion, Cynical Depersonalization, and Personal Accomplishment and Adjusted for Individual-Level Variables**

	Emotional Exhaustion			Cynical Depersonalization			Personal Accomplishment		
	BFL		BFR	BFL		BFR	BFL		BFR
	ISCED1	ISCED2	ISCED2	ISCED1	ISCED2	ISCED2	ISCED1	ISCED2	ISCED2
% students from socioeconomically disadvantaged homes (PQ)	0.01	0.05*	0.04	0.04 <sup>†</sup>	0.10***	0.05 <sup>†</sup>	-0.05*	0.02	-0.01
Intimidation or verbal abuse of teachers or staff (PQ)	0.05*	0.05*	0.03	0.04	0.09***	0.04	-0.02	-0.06**	-0.04
Aggregated collaborative school climate	-0.07**	-0.04 <sup>†</sup>	-0.08*	-0.08**	-0.03	-0.08**	0.05*	0.05**	-0.02
<i>N</i>	2,390	2,563	1,825	2,387	2,560	1,819	2,387	2,559	1,818
Schools	172	163	113	172	163	113	172	163	113

*Note.* Cell entries are standardized coefficients. School-level variables adjusted for the following individual-level variables: gender, work-experience, self-efficacy, workload, stress by intimidation, perceived teacher support, perceived collaborative school climate. BFL = Belgium Flemish-speaking community; BFR = Belgium French-speaking community; ISCED = International Standard Classification of Education; ISCED1 = primary education; ISCED2 = lower secondary education; PQ = items from principal questionnaire.

<sup>†</sup> $p < .10$ . \* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

in schools that score higher on collaborative school culture report less emotional exhaustion in primary education in the Flemish community ( $-0.07$ ;  $p < .01$ ), lower secondary education in the Flemish ( $-0.04$ ;  $p < .10$ ) and French-speaking community ( $-0.08$ ;  $p < .05$ ).

Similar relationships are found for depersonalization. The proportion of students from socioeconomically disadvantaged homes is positively related to depersonalization in the three samples, although this is only barely significant ( $0.04$ ;  $p < .10$ ) in both primary education in the Flemish community and lower secondary education in the French-speaking community. A more collaborative school climate is negatively related to individual teacher depersonalization in primary education in the Flemish community ( $-0.08$ ;  $p < .01$ ) and lower secondary education in the French-speaking community ( $-0.08$ ;  $p < .01$ ).

For the dimension of personal accomplishment, no significant relationship was found between any school-level variables and personal accomplishment in lower secondary education in the French-speaking community. In the Flemish community, a collaborative school culture was positively related to personal accomplishment in primary ( $0.05$ ;  $p < .05$ ) and lower secondary education ( $0.05$ ;  $p < .05$ ). The proportion of students from socioeconomically disadvantaged homes was negatively related to personal accomplishment, but this relationship was only statistically significant in primary education in the Flemish community ( $-0.05$ ;  $p < .05$ ). The frequency with which intimidation or verbal abuse of teachers occurs at school according to the principal was negatively related to personal accomplishment in lower secondary education in the Flemish community ( $-0.06$ ;  $p < .05$ ).

## Discussion and Conclusion

The school context is widely believed to be very important in explaining the prevalence of burnout among teachers. However,

surprisingly, little research investigates differences between schools, as most do not take the school context into account. Therefore, our aim was twofold. First, we critically reviewed the existing multilevel studies on teacher burnout that investigate school-level differences. Second, we conducted the largest empirical study on teacher burnout to date. We assessed the magnitude of between-school differences for emotional exhaustion, depersonalization, and personal accomplishment among 2,300 primary (183 schools) and 2,700 lower secondary (190 schools) teachers in the Flemish-speaking community of Belgium, and 2,135 lower secondary (120 schools) teachers in the French-speaking community of Belgium, based on the data gathered during the most recent wave of TALIS.

Our review revealed that claims about the importance of the school context in explaining teacher burnout are not as self-evident as is often presumed. One of the core problems is that conceptual confusion exists concerning school-context variables and that most studies that investigate the “school context” rely on single-level analyses. For this reason, we advise researchers to distinguish between the (1) individual-, (2) transactional- (teachers’ perception of the work environment), and (3) organizational-level (measured directly at the school level or through aggregated teacher perceptions) variables. Teacher burnout studies often confuse transactional with organizational-level variables. In such cases, single-level results are used to conclude that schools should change their leadership styles or support-environments to prevent teacher burnout. This is a classic example of an individualistic fallacy, as data at an individual level are used to formulate conclusions at the school level (Hox et al., 2017). Single-level studies are simply unable to determine how relevant the school context truly is to teacher burnout. A multilevel approach is needed to simultaneously investigate the relative contributions of individual and school-context predictors of teacher burnout.

When looking at the literature that uses a multilevel approach, it becomes clear that (1) the number of studies is limited and (2)

the observed between-school differences in teacher burnout are small. However, most studies suffer from methodological problems, such as relying on convenience samples (e.g., small non-random samples at the teacher and/or school-level that do not allow results to be generalized to the population); merging different educational levels, type of schools (e.g., general and special education schools), and types of staff (e.g., teaching assistants, heads of departments, etc.) in the analyses; using composite nonvalidated burnout measures; and not simultaneously assessing the magnitude of the general contextual (e.g., through ICC) and specific contextual effects (e.g., reporting standardized coefficients). The latter is important, as research has indicated that the likelihood of finding significant school-context effects increases as the ICC decreases (Merlo et al., 2018).

Regardless of the quality of the reviewed studies, the reported proportion of the variance that could be attributed to the school level was relatively low and ranged between 1.0% and 6.2% for emotional exhaustion; between 0.0% and 8.2% for depersonalization; and between 0.0% and 3.0% for personal accomplishment. When standardized estimates were reported at the school level, the strength of the relationships was weak. Our analyses on the TALIS 2018 data for Belgium took most of the methodological limitations of the previous studies into account. Our results indicated that the between-school variance in teacher burnout ranged between 2.1% and 5.4% for emotional exhaustion, between 3.1% and 4.0% for depersonalization, and between 0.6% and 2.0% for personal accomplishment. Moreover, we found small but consistent relationships between the three dimensions of teacher burnout and school-level job demands (e.g., the frequency of intimidation or verbal abuse of teachers occurring at school, and the proportion of students from socio-economically disadvantaged homes) and resources (e.g., collaborative school climate) school-level predictors. However, it should be underscored that such relationships are aiming to explain a small amount of between-school variance in teacher burnout.

### **Does the School Context Matter for Teacher Burnout Research?**

Substantial differences in teacher burnout between schools would provide strong indications that organizational factors at school might be important contributors to teacher burnout and that prevention strategies should be directed at the school level. Both the existing research and our analyses of three independent samples indicate that such strong evidence is simply lacking.

One interpretation of these results could be that teacher burnout is first and foremost an internal psychological phenomenon. That view is supported by the transactional model of stress and coping, which puts an individual's perception of a stressor at the center of the stress experience (Lazarus & Folkman, 1987). In other words, stress and ultimately teacher burnout may result from how individual- and school-level job demands and resources are appraised by individual teachers and this determines how they respond to the stressor. Such an interpretation clearly leaves much less room for school-specific policies to address burnout at the collective level than has often been assumed. Because few differences between schools have been

observed, little empirical evidence exists to substantiate interventions to reduce burnout by changing the way schools are organised. This implies that actions taken in this direction may have been overestimating the importance of the school context.

Specific school-level interventions might also neglect teachers in general. For example, policymakers could choose to focus interventions on schools with a high proportion of pupils from socioeconomically disadvantaged homes, as this school-level indicator is related to teacher burnout in most studies. However, because the between-school variation in teacher burnout only represents a small part of the total individual variation in teacher burnout, many teachers with high burnout levels would be ignored simply because they teach in schools with a low number of pupils from socio-economically disadvantaged homes. This illustrates the importance of simultaneously assessing the general contextual effect (e.g., the ICC) and the specific school-level predictors.

Although most variation in teacher burnout is clearly situated at the individual level, there are two reasons why the school-level variance should not be ignored. First, it can be argued that even a small preventive impact on a phenomenon as detrimental as teacher burnout is important. Indeed, our results show that fostering a collaborative school culture (e.g., creating a culture of shared responsibility for school issues) and reducing intimidation or the verbal abuse of teachers are negatively related to emotional exhaustion and depersonalization. These are aspects that school leaders can readily influence through school policies.



Second, the absence of large differences between schools might result from the measurement methods used for teacher burnout and from implicit processes of social comparison. Reference group theory holds that people compare themselves with others by using a group's values and norms as a point of reference for their own self-appraisals (Hyman & Singer, 1968). A well-known example of this is the social comparison process in the big-fish–little-pond effect, where students compare their own academic ability with that of their class or school peers and use information based on this implicit social comparison to construe their own academic self-concept (Marsh et al., 2008). The most widely used burnout measure, the MBI, uses an individual referent and could elicit similar implicit social comparison processes. For example, individual teachers could rate their own level of emotional exhaustion as high when comparing themselves with other teachers in their school but low when comparing themselves with teachers from other schools or vice versa.

Future research could explore different ways of formulating burnout items and make the comparative reference group more salient. This can be implemented by using the referent-shift consensus model, where the items are changed from an individual referent to a (comparative) group-level referent (Chan, 1998). We expect that such measures would be better to assess between-school differences. An indication for this is found in a study on perceived collective burnout by González-Morales et al. (2011) where the ICC of collective emotional exhaustion (26%) and collective depersonalization (29%) was assessed by the question stem "Teachers in this school." The between-school variance was substantially larger than the ICC of the traditionally measured scales in the same study of individual emotional exhaustion (6.2%) and depersonalization (8.2%).

Our study has several limitations. First, the cross-sectional design based on self-reported measures does not allow for causal inferences to be made between the investigated variables. Second, we chose to focus on three school-level variables. It is clear from our review that many other compositional and contextual school-level variables are potentially relevant. Although the TALIS questionnaire covers a broad range of indicators, other measures (e.g., more elaborate operationalization of interpersonal relationships) that could be relevant to teacher burnout were not included in the questionnaire. Third, variation in burnout across schools might be higher in other countries, where education systems are characterized by high segregation in terms of childhood poverty, school infrastructure, teacher quality, and so on. To assess such variation between countries, cross-national and high-quality multilevel data on teacher burnout is needed.

Notwithstanding these limitations, this study shows that although it seems self-evident and widely accepted among researchers and policymakers that the school context plays an important role in teacher burnout, existing research has mainly based its conclusions on inadequate single-level methods. We found that between-school differences for teacher burnout are small and illustrate the necessity to simultaneously assess the general and specific contextual effects. If the between-school variance in teacher burnout is generally small, then at best one is explaining a large amount of very little.

#### ORCID iDs

Filip Van Droogenbroeck  <https://orcid.org/0000-0003-1133-3495>  
Bram Spruyt  <https://orcid.org/0000-0003-0573-724X>

#### REFERENCES

References marked with an asterisk (\*) were included in the review.

Agirdag, O., & Van Houtte, M. (2011). A tale of two cities: Bridging families and schools. *Educational Leadership*, 68(8), 42–46.

Ainley, J., & Carstens, R. (2018). *Teaching and Learning International Survey (TALIS) 2018 conceptual framework* (OECD Education Working Papers No. 187). Organisation for Economic Co-operation and Development. <https://doi.org/10.1787/799337c2-en>

Ainsworth, S., & Oldfield, J. (2019). Quantifying teacher resilience: Context matters. *Teaching and Teacher Education*, 82, 117–128. <https://doi.org/10.1016/j.tate.2019.03.012>

Alarcon, G. M. (2011). A meta-analysis of burnout with job demands, resources, and attitudes. *Journal of Vocational Behavior*, 79(2), 549–562. <https://doi.org/10.1016/j.jvb.2011.03.007>

Alker, H. R., Jr. (1969). A typology of ecological fallacies. In M. Dogan & S. Rokkan (Eds.), *Quantitative ecological analysis in the social sciences* (pp. 69–86). MIT Press.

Bakker, A. B., & Demerouti, E. (2007). The job demands-resources model: State of the art. *Journal of Managerial Psychology*, 22(3), 309–328. <https://doi.org/10.1108/02683940710733115>

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Sage.

Byrne, B. M. (1994). Burnout: Testing for the validity, replication, and invariance of causal structure across elementary, intermediate, and secondary teachers. *American Educational Research Journal*, 31(3), 645–673. <https://doi.org/10.3102/00028312031003645>

Cano-García, F. J., Padilla-Muñoz, E. M., & Carrasco-Ortiz, M. A. (2005). Personality and contextual variables in teacher burnout.

*Personality and Individual Differences*, 38(4), 929–940. <https://doi.org/10.1016/j.paid.2004.06.018>

Chan, D. (1998). Functional relations among constructs in the same content domain at different levels of analysis: A typology of composition models. *Journal of Applied Psychology*, 83(2), 234–246. <https://doi.org/10.1037/0021-9010.83.2.234>

Chang, M.-L. (2009). An appraisal perspective of teacher burnout: Examining the emotional work of teachers. *Educational Psychology Review*, 21(3), 193–218. <https://doi.org/10.1007/s10648-009-9106-y>

Conley, S., & You, S. (2018). School organizational factors relating to teachers' intentions to leave: A mediator model. *Current Psychology*, 40, 379–389. <https://doi.org/10.1007/s12144-018-9953-0>

\*Ford, T. G., Olsen, J., Khojasteh, J., Ware, J., & Urick, A. (2019). The effects of leader support for teacher psychological needs on teacher burnout, commitment, and intent to leave. *Journal of Educational Administration*, 57(6), 615–634. <https://doi.org/10.1108/JEA-09-2018-0185>

\*González-Morales, M. G., Peiró, J. M., Rodríguez, I., & Bliese, P. D. (2011). Perceived collective burnout: A multilevel explanation of burnout. *Anxiety, Stress, & Coping*, 25(1), 43–61. <https://doi.org/10.1080/10615806.2010.542808>

Greenglass, E. R., Burke, R. J., & Konarski, R. (1997). The impact of social support on the development of burnout in teachers: Examination of a model. *Work & Stress*, 11(3), 267–278. <https://doi.org/10.1080/02678379708256840>

Hobfoll, S. E. (1989). Conservation of resources: A new attempt at conceptualizing stress. *American Psychologist*, 44(3), 513–524. <https://doi.org/10.1037/0003-066X.44.3.513>

Horn, J. E., & Schaufeli, W. B. (1998). *Maslach Burnout Inventory: The Dutch Educators Survey (MBI-NL-ES) psychometric evaluations. Manual* [Unpublished manuscript]. Utrecht University.

Hox, J. J., Moerbeek, M., & Schoot, R. van de. (2017). *Multilevel analysis: Techniques and applications* (3rd ed.). Routledge. <https://doi.org/10.4324/9781315650982>

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>

Hyman, H. H., & Singer, E. (1968). *Readings in reference group theory and research*. Free Press.

Jacobs, D., Rea, A., Teney, C., Callier, L., & Lothaire, S. (2009). *De sociale lift blijft steken de prestaties van allochtone leerlingen in de Vlaamse Gemeenschap en de Franse Gemeenschap*. [The social lift stops working. The performance of immigrant pupils in the Flemish community and the French community]. Koning Boudewijnstichting.

\*Kim, J., Youngs, P., & Frank, K. (2017). Burnout contagion: Is it due to early career teachers' social networks or organizational exposure? *Teaching and Teacher Education*, 66, 250–260. <https://doi.org/10.1016/j.tate.2017.04.017>

\*Klusmann, U., Kunter, M., Trautwein, U., Lüdtke, O., & Baumert, J. (2008). Engagement and emotional exhaustion in teachers: Does the school context make a difference? *Applied Psychology*, 57(s1), 127–151. <https://doi.org/10.1111/j.1464-0597.2008.00358.x>

Lai, M. H. C., & Kwok, O. (2015). Examining the rule of thumb of not using multilevel modeling: The “design effect smaller than two” rule. *Journal of Experimental Education*, 83(3), 423–438. <https://doi.org/10.1080/00220973.2014.907229>

Langballe, E. M. (2006). The factorial validity of the Maslach Burnout Inventory: General survey in representative samples of eight different occupational groups. *Journal of Career Assessment*, 14(3), 370–384. <https://doi.org/10.1177/1069072706286497>

- Lazarus, R. S., & Folkman, S. (1987). Transactional theory and research on emotions and coping. *European Journal of Personality, 1*(3), 141–169. <https://doi.org/10.1002/per.2410010304>
- Lee, R. T., & Ashforth, B. E. (1996). A meta-analytic examination of the correlates of the three dimensions of job burnout. *Journal of Applied Psychology, 81*(2), 123–133. <https://doi.org/10.1037/0021-9010.81.2.123>
- \*Lim, S., & Eo, S. (2014). The mediating roles of collective teacher efficacy in the relations of teachers' perceptions of school organizational climate to their burnout. *Teaching and Teacher Education, 44*, 138–147. <https://doi.org/10.1016/j.tate.2014.08.007>
- Lorah, J. (2018). Effect size measures for multilevel models: Definition, interpretation, and TIMSS example. *Large-Scale Assessments in Education, 6*(1), Article 8. <https://doi.org/10.1186/s40536-018-0061-2>
- Maas, C. J. M., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 1*(3), 86–92. <https://doi.org/10.1027/1614-2241.1.3.86>
- Marsh, H. W., Trautwein, U., Lüdtke, O., & Köller, O. (2008). Social comparison and big-fish-little-pond effects on self-concept and other self-belief constructs: Role of generalized and specific others. *Journal of Educational Psychology, 100*(3), 510–524. <https://doi.org/10.1037/0022-0663.100.3.510>
- Maslach, C., & Jackson, S. E. (1986). *Maslach Burnout Inventory* (2nd ed.). Consulting Psychologists Press.
- Maslach, C., & Leiter, M. P. (1999). Teacher burnout: A research agenda. In *Understanding and preventing teacher burnout: A sourcebook of international research and practice* (pp. 295–303). Cambridge University Press. <https://doi.org/10.1017/CBO9780511527784.021>
- Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job burnout. *Annual Review of Psychology, 52*(1), 397–422. <https://doi.org/10.1146/annurev.psych.52.1.397>
- \*McCarthy, C. J., Lambert, R. G., O'Donnell, M., & Melendres, L. T. (2009). The relation of elementary teachers' experience, stress, and coping resources to burnout symptoms. *Elementary School Journal, 109*(3), 282–300. <https://doi.org/10.1086/592308>
- \*McCormick, J., & Barnett, K. (2011). Teachers' attributions for stress and their relationships with burnout. *International Journal of Educational Management, 25*(3), 278–293. <https://doi.org/10.1108/09513541111120114>
- Merlo, J., Wagner, P., Austin, P. C., Subramanian, S., & Leckie, G. (2018). General and specific contextual effects in multilevel regression analyses and their paradoxical relationship: A conceptual tutorial. *SSM - Population Health, 5*, 33–37. <https://doi.org/10.1016/j.ssmph.2018.05.006>
- Musca, S. C., Kamiejski, R., Nugjer, A., Méot, A., Er-Rafiy, A., & Brauer, M. (2011). Data with hierarchical structure: Impact of intraclass correlation and sample size on type-I error. *Frontiers in Psychology, 2*, 74. <https://doi.org/10.3389/fpsyg.2011.00074>
- \*O'Brennan, L., Pas, E., & Bradshaw, C. (2017). Multilevel examination of burnout among high school staff: Importance of staff and school factors. *School Psychology Review, 46*(2), 165–176. <https://doi.org/10.17105/SPR-2015-0019.V46-2>
- Organisation for Economic Co-operation and Development. (2019). *TALIS 2018 technical report*. [https://www.oecd.org/education/talis/TALIS\\_2018\\_Technical\\_Report.pdf](https://www.oecd.org/education/talis/TALIS_2018_Technical_Report.pdf)
- Parrello, S., Ambrosetti, A., Iorio, I., & Castelli, L. (2019). School burnout, relational, and organizational factors. *Frontiers in Psychology, 10*. <https://doi.org/10.3389/fpsyg.2019.01695>
- \*Pas, E. T., Bradshaw, C. P., & Hersfeldt, P. A. (2012). Teacher and school-level predictors of teacher efficacy and burnout: Identifying potential areas for support. *Journal of School Psychology, 50*(1), 129–145. <https://doi.org/10.1016/j.jsp.2011.07.003>
- Pietarinen, J., Pyhältö, K., Soini, T., & Salmela-Aro, K. (2013). Reducing teacher burnout: A socio-contextual approach. *Teaching and Teacher Education, 35*, 62–72. <https://doi.org/10.1016/j.tate.2013.05.003>
- \*Ross, S. W., Romer, N., & Horner, R. H. (2012). Teacher well-being and the implementation of school-wide positive behavior interventions and supports. *Journal of Positive Behavior Interventions, 14*(2), 118–128. <https://doi.org/10.1177/1098300711413820>
- Sampson, R. J., & Groves, W. B. (1989). Community structure and crime: Testing social-disorganization theory. *American Journal of Sociology, 94*(4), 774–802. <https://doi.org/10.1086/229068>
- Schaufeli, W. B., Bakker, A. B., Hoogduin, K., Schaap, C., & Kladler, A. (2001). On the clinical validity of the maslach burnout inventory and the burnout measure. *Psychology & Health, 16*(5), 565–582. <https://doi.org/10.1080/08870440108405527>
- Schaufeli, W. B., Daamen, J., & Van Mierlo, H. (1994). Burnout among Dutch teachers: An Mbi-validity study. *Educational and Psychological Measurement, 54*(3), 803–812. <https://doi.org/10.1177/0013164494054003027>
- Schaufeli, W. B., & Enzmann, D. (1998). *The burnout companion to study and practice: A critical analysis*. Taylor & Francis.
- \*Shackleton, N., Bonell, C., Jamal, F., Allen, E., Mathiot, A., Elbourne, D., & Viner, R. (2019). Teacher burnout and contextual elements of school environment. *Journal of School Health, 89*(12), 977–993. <https://doi.org/10.1111/josh.12839>
- Skaalvik, E. M., & Skaalvik, S. (2009). Does school context matter? Relations with teacher burnout and job satisfaction. *Teaching and Teacher Education, 25*(3), 518–524. <https://doi.org/10.1016/j.tate.2008.12.006>
- Skaalvik, E. M., & Skaalvik, S. (2017). Still motivated to teach? A study of school context variables, stress and job satisfaction among teachers in senior high school. *Social Psychology of Education, 20*(1), 15–37. <https://doi.org/10.1007/s11218-016-9363-9>
- Statacorp. (2019). *Stata statistical software: Release 16*.
- Steinhardt, M. A., Jaggars, S. E. S., Faulk, K. E., & Gloria, C. T. (2011). Chronic work stress and depressive symptoms: Assessing the mediating role of teacher burnout. *Stress Health, 27*(5), 420–429. <https://doi.org/10.1002/smi.1394>
- Subramanian, S. V., Jones, K., Kaddour, A., & Krieger, N. (2009). Revisiting Robins: The perils of individualistic and ecologic fallacy. *International Journal of Epidemiology, 38*(2), 342–360. <https://doi.org/10.1093/ije/dyn359>
- \*Ullrich, A., Lambert, R. G., & McCarthy, C. J. (2012). Relationship of German elementary teachers' occupational experience, stress, and coping resources to burnout symptoms. *International Journal of Stress Management, 19*(4), 333–342. <https://doi.org/10.1037/a0030121>
- Vandenberghe, R., & Huberman, A. M. (1999). *Understanding and preventing teacher burnout: A sourcebook of international research and practice*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511527784>
- Van Droogenbroeck, F., Spruyt, B., & Vanroelen, C. (2014). Burnout among senior teachers: Investigating the role of workload and interpersonal relationships at work. *Teaching and Teacher Education, 43*, 99–109. <https://doi.org/10.1016/j.tate.2014.07.005>
- \*Van Maele, D., & Van Houtte, M. (2015). Trust in school: A pathway to inhibit teacher burnout? *Journal of Educational Administration, 53*(1), 93–115. <https://doi.org/10.1108/JEA-02-2014-0018>
- Vercambre, M.-N., Brosselin, P., Gilbert, F., Nerrière, E., & Kovess-Masféty, V. (2009). Individual and contextual covariates of burnout: A cross-sectional nationwide study of French teachers. *BMC Public Health, 9*(1), Article 333. <https://doi.org/10.1186/1471-2458-9-333>

\*Zheng, Q., Li, L., Chen, H., & Loeb, S. (2017). What aspects of principal leadership are most highly correlated with school outcomes in China? *Educational Administration Quarterly*, 53(3), 409–447. <https://doi.org/10.1177/0013161X17706152>

#### AUTHORS

**FILIP VAN DROOGENBROECK**, PhD, is a research professor of sociology at Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium; [filip.van.droogenbroeck@vub.be](mailto:filip.van.droogenbroeck@vub.be). His research focuses on educational and youth sociology.

**BRAM SPRUYT**, PhD, is an associate professor of sociology at Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium; [bram.spruyt@vub.be](mailto:bram.spruyt@vub.be). His research focuses on teachers, social inequality, and public opinion research.

**VALÉRIE QUITTRE**, MSc, biochemistry, is a senior researcher in educational science at the University of Liège, Place des Orateurs, 2, B32, 4000 Liège, Belgium; [v.quittre@uliege.be](mailto:v.quittre@uliege.be). Her research focuses on science teaching and international large-scale assessments.

**LAFONTAINE DOMINIQUE**, PhD, is a full professor of educational science at the University of Liège, Place des Orateurs, 2, B32, 4000 Liège, Belgium; [dlafontaine@uliege.be](mailto:dlafontaine@uliege.be). Her research focuses on international large-scale assessments and educational policies.

Manuscript received April 6, 2020

Revisions received September 23, 2020,

December 14, 2020, and December 24, 2020

Accepted December 31, 2020