

1 Running head: STUDENT-CONTINGENT TEACHER EVALUATION

2

3

4

5

6

7

8

9

10 Does Teacher Evaluation based on Student Performance Predict Motivation, Well-

11 Being, and Ill-Being?

12

13

14

15

16

17

18

19

20

21

## 1 Abstract

2 This study tests an explanatory model based on self-determination theory, which posits  
3 that pressure experienced by teachers when they are evaluated based on their students'  
4 academic performance will differentially predict teacher adaptive and maladaptive  
5 motivation, well-being, and ill-being. A total of 360 Spanish physical education  
6 teachers completed a multi-scale inventory. We found support for a structural equation  
7 model that showed that perceived pressure predicted teacher autonomous motivation  
8 negatively, predicted amotivation positively, and was unrelated to controlled  
9 motivation. In addition, autonomous motivation predicted vitality positively and  
10 exhaustion negatively, whereas controlled motivation and amotivation predicted vitality  
11 negatively and exhaustion positively. Amotivation significantly mediated the relation  
12 between pressure and vitality and between pressure and exhaustion. The results  
13 underline the potential negative impact of pressure felt by teachers due to this type of  
14 evaluation on teacher motivation and psychological health.

15 *Keywords:* Self-determination; vitality; burnout; mediation; structural equation  
16 model.

17

18

19

20

21

22

23

1 Does Teacher Evaluation based on Student Performance Predict Motivation, Well-  
2 Being, and Ill-Being?

3 The Teaching and Learning International Survey (TALIS) of the Organization  
4 for Economic Co-operation and Development (OECD, 2014) notes that the most  
5 important factor for quality education is teacher performance. Teachers are the main  
6 agents for engaging students in school tasks and promoting their learning (Rockoff,  
7 2004). In this regard, numerous studies have examined ways in which to measure  
8 teacher performance (e.g., Pas, Bradshaw, & Hershfeldt, 2012). Traditionally, there are  
9 two elements involved in evaluating teacher performance: supervision (the formative  
10 aspect) and evaluation (the summative aspect). Supervision involves the assessment of  
11 teachers' lesson plans, their teaching skills and instructional strategies, and how well  
12 they have mastered the material. On the other hand, the summative element involves  
13 evaluating how well the students have learned the lesson content delivered by the  
14 teacher based on the students' performance on assessments or their grades.

15 In this paper, we focus on teacher evaluation that is based on student  
16 performance. According to the Teaching and Learning International Survey (TALIS)  
17 report (Ministry of Education, Culture and Sport, 2014), the most widely used  
18 procedure for teacher evaluation across several countries (e.g., United Kingdom,  
19 Sweden, France) of the Organization for Economic Co-operation and Development  
20 (OECD) is based on students' academic grades. In Spain, for example, where current  
21 legislation stipulates the need to evaluate the performance of teachers (Marina, Pellicer,  
22 & Manso, 2015), the most widely used procedure for teacher evaluation (used in 97% of  
23 schools) is based on students' academic grades. The respective 'weight' or importance  
24 of students' grades in the evaluation of teachers' performance can, however, vary from  
25 one school to another and students' grades may also be used alongside other forms of

1 evaluation in Spain [e.g., classroom observation (59% of schools), student surveys  
2 (72%), or knowledge assessment (34%) (OECD, 2014)]. Principals who conduct formal  
3 evaluations of their teachers reported that their evaluations could affect career progress,  
4 changes in work responsibilities or, in some extreme cases, the dismissal of teachers  
5 (OECD, 2014). Perhaps not surprisingly, most of the teachers questioned in the TALIS  
6 reported disagreement with the current evaluation system and with the feedback they  
7 receive through this process (Ministry of Education, Culture and Sport, 2014). Given  
8 that the use of teacher evaluation based on students' performance has spread  
9 considerably in various countries in recent years (Isore, 2009), research on the impacts  
10 of this type of evaluation is important.

11         Despite its increasing use in various countries, some authors have suggested that  
12 external incentives are not always effective in improving teaching performance. For  
13 example, Yuan et al. (2014) found that incentive pay programs did not improve the  
14 practices and motivation of teachers. Furthermore, in a review article, Firestone (2014)  
15 indicated that incentives programs that use performance-based pay in educational  
16 contexts to improve student performance are ineffective and can undermine the intrinsic  
17 incentives of the teachers.

18         The practice of providing external incentives to teachers that are contingent on  
19 their students' performance is linked to growth models for evaluation in education. Such  
20 models aspire to measure the specific contribution of teachers to the growth of their  
21 students (McCaffrey, Lockwood, Koretz, & Hamilton, 2003). Value-added models are  
22 some of the most widely used types of growth models. Specifically, these models try to  
23 capture student performance over time (i.e., the development of knowledge or skills) as  
24 a consequence of student experiences in schools (Harvey, 2004). Despite their  
25 widespread use, the effectiveness of value-added models has been questioned on a

1 number of grounds. For example, research has shown that school factors (including  
2 teacher performance) account for only about 20% of the variance in student  
3 performance (Berliner, 2014). In addition, Rothstein (2010) identified several additional  
4 external factors that can influence student performance, including students' experiences  
5 with previous teachers, the simultaneous influence of different teachers, the number of  
6 students in the class, the inclusion of students with special educational needs,  
7 curriculum materials, and the sociodemographic characteristics of the school and its  
8 students.

9         Given the number of different factors that may interact to influence student  
10 performance — many of which are outside of the teacher's control — it has been argued  
11 that student performance outcomes do not appear to be a sufficiently robust means by  
12 which to assess teacher effectiveness or to warrant consideration in decisions that may  
13 affect a teacher's career (McCaffrey, Sass, Lockwood, & Kata, 2009). In fact, various  
14 negative consequences resulting from the use of this type of evaluation have recently  
15 been documented, including a compression of the curriculum, decreased collaborative  
16 work between educators, and discouragement of teachers to work with the neediest  
17 students (Baker et al., 2010; Hewitt, 2015). In addition, there is growing evidence to  
18 indicate that the pressure felt by teachers as a result of this form of evaluation may also  
19 have negative repercussions for their psychological health.

## 20 **Teacher Evaluation and Psychological Health**

21         The limited research examining the impact of teacher evaluation based on  
22 student performance has shown that teachers experience increased stress, pressure, and  
23 anxiety as a consequence of such evaluations (Hewitt, 2015; Jiang, Spote, & Luppescu,  
24 2015; von der Embse, Pendergast, Segool, Saeki, & Ryan, 2016). For example,  
25 Goldhaber and Hannaway (2004) found that evaluation-related pressure and anxiety

1 levels were high not only among teachers in the U.S. whose schools had poor results  
2 and who attempted to improve those results, but also among teachers whose schools  
3 exhibited high performance and who tried to maintain that high level. Furthermore,  
4 accountability for student outcomes has been shown to be associated with increased  
5 teacher anxiety, and decreased teacher motivation, particularly among teachers who do  
6 not achieve the objectives set by the administration, irrespective of how much they have  
7 endeavored to do so (Feng, Figlio, & Sass 2010; Finnigan & Gross, 2007).

8         Despite increasing interest in the topic, the specific impact of student  
9 performance-based teacher evaluation on the well-being and ill-being of teachers needs  
10 more empirical attention in an effort to explore mechanisms that could mediate such an  
11 impact. (Taylor & Tyler, 2012). To this end, Self-Determination Theory (SDT; Deci &  
12 Ryan, 1985) and its focus on motivation- may prove as a helpful conceptual framework.

### 13 **Self-Determination Theory**

14         Various studies (e.g., Cuevas, Sanchez-Oliva, Bartholomew, Ntoumanis, &  
15 Garcia-Calvo, 2015; Taylor & Ntoumanis, 2007; Taylor, Ntoumanis, & Standage, 2008)  
16 have noted the usefulness of Self-Determination Theory (SDT; Deci & Ryan, 1985) for  
17 the study of teacher motivation and psychological health. SDT is a widely applied  
18 theoretical approach to the study of human motivation, development, and well-being.  
19 The theory focuses on different ‘types’ of motivation which have been shown to predict  
20 a diverse range of adaptive and maladaptive cognitive, affective, and behavioral  
21 outcomes. Specifically, Deci and Ryan (1985) differentiated between three different  
22 forms of motivation. First, autonomous motivation is volitional and reflects interest or  
23 personal value. For example, when an activity is performed for pleasure or personal  
24 growth. Second, controlled motivation reflects external and/or internal contingencies  
25 and pressures. For example, when an activity is performed for external incentives, such

1 as money or social recognition. Finally, amotivation reflects a lack of both intrinsic and  
2 extrinsic motivation. Individuals engage passively in activities without any sense of  
3 intention. Several studies have linked teacher motivation with teacher well-being or ill-  
4 being. For instance, higher scores on autonomous motivation have been shown to be  
5 positively associated with higher levels of well-being, and negatively associated with  
6 higher levels of ill-being in Israeli teachers (Roth, Assor, Kanat-Maymon, & Kaplan,  
7 2007). In the same country, Eyal and Roth (2011) found that burnout in teachers was  
8 negatively predicted by autonomous motivation and positively predicted by controlled  
9 motivation. In addition, Fernet, Guay, Senécal, & Austin (2012) found that autonomous  
10 motivation negatively predicted emotional exhaustion in French-Canadian teachers.

11         Although the association between teacher motivation and well-being/ill-being  
12 has been explored within the literature, the influence of teacher evaluations on both  
13 adaptive and maladaptive types of teacher motivation and their psychological health is  
14 less well understood. Given that a lack of personal control has been associated with ill-  
15 being (Weiner, 2004), it may be particularly important to consider whether the link  
16 between student performance-contingent evaluations and teachers' psychological health  
17 is related to the perceived lack of control, or self-determination, that teachers perceive  
18 in relation to their students' performance (Berliner, 2014; Konstantopoulos, 2014).  
19 Specifically, evaluations based on external and largely non-controllable criteria, such as  
20 student performance, are likely to be perceived as controlling and hence have the  
21 potential to undermine self-determined motivation (Deci & Ryan, 1985). In such  
22 situations, teachers are likely to feel controlled in their motivation to work or even  
23 amotivated. In turn, such motivational states are unlikely to be conducive to the  
24 nurturing of one's well-being. Despite this, the role of teacher motivation as a mediator

1 between student performance-contingent teacher evaluation and well-being and ill-being  
2 has not been explored.

### 3 **The Present Study**

4 In this study, we propose and test a model, using data collected from physical  
5 education teachers, that links perceived pressure due to evaluations dependent on  
6 student performance, with teacher motivation and, in turn, well-being and ill-being.

7 Whilst we acknowledge that teacher evaluation consists of multiple dimensions (Isore,  
8 2009), we are specifically interested in this particular type of evaluation because it is  
9 becoming increasingly used in educational settings and has the potential to undermine  
10 the psychological health of teachers. Specifically, the objectives of the study were to (a)  
11 analyze the association between the perceived pressure caused by teacher evaluation  
12 based on student performance and teacher psychological well-being and ill-being and  
13 (b) examine the different types of teacher motivation, as outlined by SDT, as possible  
14 mediators of this association. To this end, four hypotheses were proposed. First, based  
15 on previous research that has indicated that teacher evaluations influence teacher well-  
16 being and ill-being (e.g., Dworkin & Tobe, 2014; Hewitt, 2015; von der Embse et al.,  
17 2016), we hypothesized that perceived pressure associated with teacher evaluation  
18 based on student performance would have a direct and negative effect on vitality and a  
19 direct and positive effect on exhaustion (H1). Second, based on SDT and previous  
20 research that has shown that controlling environments can undermine motivation (e.g.,  
21 Finnigan & Gross 2007; Yuan et al., 2012), it was hypothesized that perceived pressure  
22 related to student performance evaluation would negatively predict autonomous  
23 motivation and positively predict controlled motivation and amotivation (H2). Third, in  
24 line with SDT and previous research (e.g., Eyal & Roch, 2011; Fernet et al., 2012), we  
25 anticipated that autonomous motivation would positively predict vitality (an indicator of

1 well-being) and negatively predict exhaustion (an indicator of ill-being), whereas  
2 controlled motivation and amotivation would negatively predict vitality and positively  
3 predict exhaustion (H3). Finally, we hypothesized that autonomous and controlled  
4 motivation, as well as amotivation, would significantly mediate the relations between  
5 pressure due to student performance-based evaluation and teacher vitality and  
6 exhaustion (H4; e.g., Berliner, 2014; Konstantopoulos, 2014).

## 7 **Method**

### 8 **Participants and Educational Context**

9 A total of 360 Caucasian physical education teachers of secondary education  
10 (230 men and 130 women) from all regions of Spain participated in the study. The  
11 participants were between 23 and 61 years of age ( $M = 40.51$ ;  $SD = 9.01$ ), their work  
12 experience ranged between 1 and 39 years ( $M = 14.57$  years,  $SD = 9.62$  years), and they  
13 were employed in public ( $n = 335$ ) and private ( $n = 25$ ) schools. In the Spanish  
14 educational system, physical education is a mandatory subject and it has similar  
15 curricular structure, academic goals, and assessment systems to those of other subjects  
16 (Pastor, Brunicardi, Arribas, & Aguado, 2016). Further, physical education teachers  
17 have the same staff development and promotion opportunities as other teachers.  
18 Alongside teaching a number of sports and games, the physical education curriculum  
19 requires the teaching of anatomy and physiology related to the study of general health  
20 and physical fitness.

21 The Spanish educational system has a curricular model focused on competency  
22 development. That is, all subjects, including physical education, should contribute to the  
23 development of key competencies (motor, linguistic, mathematics, digital, social,  
24 cultural, learning to learn, or entrepreneurial initiative). Students are evaluated through  
25 different methods, such as practical tests, teacher observations, written examinations

1 and homework assignments. Through these methods, the teacher evaluates and grades  
2 the students using the criteria and standards established by the Ministry of Education. In  
3 addition, the physical education grades received by the students contribute to their  
4 global academic record (Organic Law for the Improvement of the Educative Quality,  
5 2013).

## 6 **Procedure**

7 Schools and professional associations of physical education teachers were  
8 contacted and informed about the research objectives. These institutions approved and  
9 supported the project and facilitated contact with the participants via email. Following  
10 ethics approval from a Spanish university, consent from all participants was obtained.  
11 Participants were informed that the study was voluntary and their responses would be  
12 kept anonymous. The questionnaire was completed online and the measures were  
13 completed in the same order by all participants.

## 14 **Measures**

15 **Perceived pressure due to student performance.** A subscale of the Pressure at  
16 Work scale (Pelletier, Seguin-Levesque, & Legault, 2002; Taylor et al., 2008), adapted  
17 to the Spanish context (Bartholomew, Ntoumanis, Cuevas, & Lonsdale, 2014), was used  
18 to assess the pressure perceived by teachers for their students to display strong academic  
19 performance. The Pressure for Evaluation based on Student Performance subscale  
20 consists of four items that measure the pressure perceived by the teacher when being  
21 evaluated according to the performance of his or her students (e.g., “My school will  
22 evaluate me poorly if my students don’t get good grades”). The response range was 1  
23 (*not at all true*) to 7 (*very true*). The scale has been used previously with physical  
24 education teachers in both Spanish (Bartholomew et al., 2014) and English (Taylor et  
25 al., 2008), with evidence of adequate reliability (Cronbach alphas of .79 and .75,

1 respectively) and validity ( $\chi^2 (98) = 268.3, p < .001, CFI = .91, TLI = .90, RMSEA =$   
2  $.07$ ; Bartholomew et al., 2014).

3       **Motivation.** The Work Motivation Inventory (Blais, Lachance, Vallerand,  
4 Briere, & Riddle, 1993), adapted to the Spanish educational context (Cuevas, Sánchez-  
5 Oliva, Contreras, Moreno, & García-Calvo, 2014), was used to measure the different  
6 types of teacher motivation outlined within SDT (Deci & Ryan, 1985). The heading  
7 “Why do you teach?” was followed by six four-item subscales tapping each regulation:  
8 intrinsic motivation (e.g., “For the intense moments of pleasure teaching gives me”),  
9 integrated motivation (e.g., “Teaching is part of my life”), identified motivation (e.g., “I  
10 want to pursue my career in teaching”), introjected motivation (e.g., “I want to succeed  
11 at teaching, if not I would be very ashamed of myself”), external motivation (e.g., “For  
12 the income it provides me”), and amotivation (e.g., “I don’t know, I have the impression  
13 that I don’t have what it takes to teach”). The response scale used ranged from 1 (*not at*  
14 *all true*) to 7 (*very true*) scale. Following the SDT conceptualization (Deci & Ryan,  
15 1985), the scores from the intrinsic, integrated and identified motivation subscales were  
16 averaged to form a single dimension termed “autonomous motivation”. In addition, the  
17 introjected and external motivation subscales were averaged to form a single variable  
18 termed “controlled motivation”. This combination of subscales has been previously  
19 used in a number of studies with strong psychometric evidence (e.g., autonomous  
20 motivation  $\alpha = .85$ , controlled motivation  $\alpha = .76$ ; Vansteenkiste, Lens, De Witte, De  
21 Witte, & Deci, 2004). In addition, Cuevas et al. (2014) found adequate reliability  
22 ( $\alpha > .76$  for the intrinsic, controlled and amotivation subscales) and validity ( $\chi^2 (234) =$   
23  $780.91, p < 0.01, CFI = 0.93, TLI = 0.96, RMSEA = 0.08$ ) for the Spanish version of  
24 the scale.

1           **Vitality.** The Spanish adaptation (Balaguer, Castillo, Alvarez, & Duda, 2005) of  
2 the Subjective Vitality Scale (SVS; Ryan and Frederick, 1997) was used to measure the  
3 feeling of being full of energy and alive. The instrument consists of six items (e.g., “I  
4 feel alive and full of vitality”) that are assessed using a 1 (*not at all true*) to 7 (*very true*)  
5 scale. Previous studies have reported adequate factorial structure ( $\chi^2(8) = 19.95, p <$   
6  $0.01$ , CFI = 0.97, NFI = 0.95, RMSEA = 0.08; Bostic, Rubio & Hood, 2000) for the  
7 English version, and adequate reliability ( $\alpha > .84$ ; Álvarez, Balaguer, Castillo & Duda,  
8 2012; Bostic et al., 2000) for the Spanish and English versions of the scale.

9           **Exhaustion.** The Spanish adaptation of the Exhaustion subscale (Gil-Monte,  
10 2002) within the reduced version of the Maslach Burnout Inventory (MBI; Schaufeli  
11 Leiter, Maslach, & Jackson, 1996) was used to assess mental exhaustion among the  
12 participants. It consists of five items designed to assess the mental fatigue and the  
13 decreased emotional resources of the participants (e.g., “Because of my job, I am  
14 exhausted”). The response range was 1 (*not at all true*) to 7 (*very true*). Gil-Monte  
15 (2002) and Bartholomew et al. (2014) reported evidence that supported the reliability  
16 ( $\alpha = .84$  and  $.86$ , respectively) and validity ( $\chi^2(101) = 333.17, p < 0.01$ , CFI = 0.92,  
17 TLI = 0.91, RMSEA = 0.08; Bartholomew et al., 2014) of this scale in Spanish  
18 populations.

### 19 **Data analysis**

20           Preliminary analyses were conducted using IBM-SPSS 20.0. In addition to  
21 Cronbach’s alpha, composite reliability (CR) and average variance extracted (AVE)  
22 were calculated. CR indicates the degree of consistency of the observed variables with  
23 the measurement latent construct. AVE indicates the variance of the items captured by  
24 the latent construct compared to the variance captured by measurement error. Hair,  
25 Black, Babin, and Anderson (2010) considered acceptable values to be if CR is higher

1 or equal to 0.07 and if AVE is higher or equal to 0.05. Means, standard deviations and  
2 bivariate correlations were also estimated.

3       Confirmatory factor analysis (CFA) and structural equation modeling were  
4 performed using AMOS 18.0. Factorial validity was tested with a CFA of the  
5 measurement model. For the structural equation modeling analysis, latent factors that  
6 correspond to pressure, amotivation, vitality, and exhaustion were estimated using the  
7 items from each scale as indicators. The latent factor for autonomous motivation was  
8 estimated based on the average values of intrinsic motivation and integrated and  
9 identified regulation. The latent factor for controlled motivation was estimated based on  
10 the averages of introjected and external regulation. Due to lack of normality in the data,  
11 the maximum-likelihood estimation method with bootstrapping was used. Bootstrapping  
12 provides robust standard errors estimates in the absence of normality (Byrne, 2001).  
13 The following indices were used to interpret model fit: the chi-square value, the  
14 comparative fit index (CFI), the Tucker-Lewis index (TLI), and the root mean square  
15 error of approximation (RMSEA). A model may be considered to be acceptable if CFI  
16 and TLI are close to or exceed 0.95 and if RMSEA is less than or equal to 0.08 (Hooper,  
17 Coughlan, & Mullen, 2008; Hu & Bentler, 1999; Tabachnick & Fidell, 2007).  
18 Additionally, with the RMSEA value, a confidence interval (i.e., 90%) is generated to  
19 indicate the level of the RMSEA precision. Quintana and Maxwell (1999) consider  
20 model fit to be adequate if the upper limit of this confidence interval is below 0.08 and  
21 if the range of the interval is smaller than 0.05.

22       Finally, direct and indirect effects (mediation analysis) were calculated using the  
23 bootstrapping technique proposed by Preacher and Hayes (2008) via IBM-SPSS 20.0.  
24 Bootstrapping generates a confidence interval (e.g., 95%) for indirect effects; if zero is

1 included in the confidence interval, then the indirect effect is considered to be non-  
2 significant.

### 3 **Results**

#### 4 **Preliminary Analyses**

5 The CFA results supported the validity of the measurement model: ( $\chi^2$  (150) =  
6 259.63,  $p < 0.01$ , CFI = 0.97, TLI = 0.96, RMSEA (90% CI) = .049 (.040-0.058). The  
7 means, standard deviations, Cronbach's alphas, composite reliabilities, and average  
8 mean extracted for each factor are presented in Table 1. On average, participants  
9 reported low mean levels of perceived pressure (2.78), amotivation (2.38), and  
10 exhaustion (2.75), whereas they reported high levels of autonomous motivation (5.31)  
11 and vitality (5.55). The reliability estimates were satisfactory (i.e.  $> .70$ ) for all of the  
12 variables except amotivation, for which Cronbach's alpha (.67) and composite  
13 reliability (.65) were marginally acceptable (Hair et al., 2010). Table 1 also includes  
14 bivariate correlations, which were mostly in line with theoretical predictions in that  
15 perceived pressure from evaluations based on student performance was positively  
16 associated with controlled motivation (a weak association), amotivation (a moderate  
17 association), and exhaustion (a moderate association), and was negatively but relatively  
18 weakly associated with vitality. Moreover, amotivation was positively and relatively  
19 strongly associated with exhaustion and negatively and moderately associated with  
20 vitality whilst opposite relations were observed between autonomous motivation and  
21 exhaustion (i.e., a moderate negative association) and vitality (i.e., a moderate positive  
22 association).

#### 23 **Structural Equation Model**

24 The hypothesized model (Figure 1) assumed that pressure due to evaluation  
25 would negatively predict autonomous motivation, whereas it would positively predict

1 controlled motivation and amotivation. In addition, autonomous motivation would  
2 positively predict vitality and negatively predict exhaustion, whereas controlled  
3 motivation and amotivation would positively predict exhaustion and negatively predict  
4 vitality. The model demonstrated acceptable fit indices ( $\chi^2(168) = 364.19, p < 0.01,$   
5  $CFI = 0.95, TLI = 0.94, RMSEA = 0.06$  (90% CI = 0.05-0.06). All of the hypothesized  
6 relations were significant ( $p < .01$ ) except that between perceived pressure and  
7 controlled motivation, which was not significant (providing partial support for H2 &  
8 H3).

### 9 **Direct and Indirect Effects**

10 Table 2 presents the direct effects of pressure due to evaluation on vitality and  
11 exhaustion, and the indirect effects of pressure on vitality and exhaustion through the  
12 motivation variables. Perceived pressure directly and positively predicted exhaustion  
13 and negatively predicted vitality (providing support of H1). The total indirect effects  
14 were significant. When examining the specific indirect effects, it was observed that  
15 these effects were not significant for autonomous and controlled motivation. In contrast,  
16 amotivation mediated the negative indirect effect from pressure on vitality, and the  
17 positive indirect effect of pressure on exhaustion (providing partial support for H4).

### 18 **Discussion**

19 The primary purpose of the current study was to test a model, based on SDT,  
20 that examined whether the pressure experienced by physical education teachers when  
21 evaluated based on the performance of their students was related to different types of  
22 motivation for teaching and, in turn, to teacher psychological well-being (vitality) and  
23 ill-being (exhaustion). The mediating role of autonomous motivation, controlled  
24 motivation, and amotivation in the relationship between perceived pressure and both  
25 vitality and exhaustion was also tested. There is a dearth of studies linking teacher

1 evaluations based on student performance with teacher motivation, well-being, and ill-  
2 being. Hence, this study offers important empirical evidence regarding the effects of this  
3 type of teacher evaluation.

4         The first hypothesis that perceived pressure related to teacher evaluation based  
5 on student performance would negatively predict vitality and positively predict  
6 exhaustion was fully supported. The direct effects of perceived pressure on vitality and  
7 exhaustion were significant and in the expected direction. These results were consistent  
8 and comparable in size with previous findings that associated the pressure due to this  
9 type of teacher evaluation with stress (von der Embse, 2016), anxiety (Goldhaber &  
10 Hannaway, 2004; Hewitt, 2015), and burnout (Dworkin & Tobe, 2014) among teachers.  
11 Given that teacher well-being has been linked with better teaching performance  
12 (Klusmann, Kunter, Trautwein, Lüdtke, & Baumert, 2008; Roth et al., 2007), whereas  
13 teacher ill-being has been linked to negative teacher-student relationships (von der  
14 Embse et al., 2016) and poorer student academic performance (Blandford, 2000), these  
15 findings suggest that evaluations based on student performance should be implemented  
16 very cautiously.

17         The second hypothesis, which proposed that perceived pressure related to  
18 student performance evaluation would negatively predict autonomous motivation and  
19 positively predict controlled motivation and amotivation, was partially supported.  
20 Perceived pressure negatively predicted teacher autonomous motivation. That is, the  
21 greater pressure that teachers felt from performance-based evaluations, the less likely  
22 they were to report that they taught for reasons of interest and personal value. These  
23 results align with those of previous studies that have reported a negative association  
24 between pressure due to teaching assessment and internal (i.e., more self-determined)  
25 forms of motivation (Finnigan & Gross 2007; Yuan et al., 2012). In addition, we found

1 that pressure due to teacher evaluation based on student performance also positively  
2 predicted amotivation among teachers. In other words, the greater pressure that teachers  
3 felt from performance-based evaluations, the more likely they were to report a complete  
4 absence of motivation to teach (i.e., their decision to teach was neither intrinsically nor  
5 extrinsically motivated). Taken together, these findings suggest that perceived pressure  
6 due to teacher evaluations based on student performance could undermine autonomous  
7 motivation and promote teacher amotivation. Motivational deficits could promote  
8 teacher ill-being (Eyal & Roth, 2011) which, in turn, could adversely affect the quality  
9 of the teacher's professional work (Klusmann et al., 2008).

10         However, it must also be noted that, contrary to what was hypothesized,  
11 pressures associated with this type of student-based teacher evaluation did not  
12 significantly predict controlled motivation in teachers. It is likely that the threats  
13 associated with the punitive aspects of this method of evaluation were perceived to be  
14 far greater than any rewards that would be available from improved student  
15 performance (Marina et al., 2015). For example, the consequences of poor teacher  
16 evaluations in Spain include difficulties in career progression, a reduction in  
17 responsibilities and status, and even job loss (Ministry of Education, Culture and Sport,  
18 2014). However, the scale assessing external regulation (one of the two components of  
19 controlled motivation) used in the present study mainly captures positive rewards (e.g.,  
20 economic gains and job security), which could explain the absence of relation between  
21 perceived pressure and controlled motivation. Hence, future studies in this area should  
22 ensure that the assessment of external regulation captures both punishments and  
23 rewards. Introjected regulation refers to internal pressures, hence, it is less likely that  
24 this component of controlled motivation would be predicted by external pressures  
25 associated with student-based teacher evaluation.

1           The third hypothesis, which concerned the relations between motivation, well-  
2 being, and ill-being, was fully supported. In line with SDT, autonomous motivation  
3 positively predicted vitality and negatively predicted exhaustion. That is, participants  
4 who reported that they taught for pleasure or personal growth were more likely to report  
5 feeling full of life and less likely to report feeling exhausted. These findings further  
6 underscore the importance of motivation based on autonomous factors, such as interest  
7 or value, for nurturing teacher well-being and psychological health (Eyal & Roth, 2011;  
8 Fernet et al., 2012; Roth et al., 2007). This is important because autonomously  
9 motivated teachers facilitate supportive teaching environments which, in turn, promote  
10 students' autonomous motivation for learning (Roth et al., 2007). On the other hand,  
11 controlled motivation and amotivation were negatively associated with vitality and  
12 positively associated with exhaustion. These results are also in line with other findings  
13 (Firestone, 2014, Yuan et al., 2012), which have indicated the limited effectiveness of  
14 initiatives aimed at improving teacher motivation based on extrinsic factors (e.g., higher  
15 payment for teachers who achieve better student performance) without considering the  
16 internal motivational resources of teachers.

17           The fourth hypothesis of our study posited that the different types of motivation  
18 would play a mediating role between perceived pressure due to student performance and  
19 teacher well-being and ill-being. This hypothesis was only partially supported.  
20 Specifically, the total indirect effects of perceived pressure on vitality and exhaustion  
21 were significant. However, an analysis of the specific indirect effects indicated that only  
22 amotivation (which, in contrast to autonomous and controlled motivation, represents the  
23 complete lack of motivation) played a mediating role in the relations between pressure  
24 and vitality, and between pressure and exhaustion. In other words, perceived pressure  
25 resulted in higher levels of exhaustion and lower levels of vitality via increased

1 amotivation. These results indicate that mediation only occurs in the absence of  
2 motivation (amotivation). That is to say, perceived pressure predicts low vitality and  
3 high exhaustion because it makes teachers experience a sense of helplessness. It is  
4 possible that controlled motivation was not a significant mediator because the  
5 questionnaire assessing the external regulation component of controlled motivation  
6 focused primarily on the rewards aspect of control (e.g., financial gains). It would have  
7 been beneficial to also assess the punishment aspect of external regulation (e.g.,  
8 penalties or delays in career progression), which is more likely to be associated with this  
9 type of teacher evaluation. In addition, from a statistical perspective, controlled and  
10 autonomous motivation may not have been significant mediators because the direct  
11 effects from perceived pressure to amotivation and, in turn, from amotivation to the two  
12 dependent variables were very strong, leaving little unique variance for autonomous and  
13 controlled motivation to account for. Such findings are particularly useful because they  
14 contribute to the understanding of potential mechanisms through which pressure due to  
15 student performance-contingent teacher evaluation can affect teacher well-being and ill-  
16 being.

### 17 **Limitations and Directions for Future Research**

18         This study has a number of limitations that should be considered in future  
19 research. First perceived pressure from student performance-contingent evaluation was  
20 fairly low in the current sample ( $M = 2.78$ ,  $SD = 0.92$ ), perhaps due to the subject that  
21 these teachers were teaching. The observed results may be different in a sample in  
22 which these pressures were felt more intensely by teachers. Second, the sample of  
23 teachers was confined to a single country, a single subject, and a single academic level.  
24 Therefore, the generalizability of the current findings to other countries, school subjects,  
25 and academic levels is unknown and should be explored. Third, the study was cross-

1 sectional in nature, which prevented us from testing causal relations between variables.  
2 Consequently, new experimental studies could complement the results of our study by  
3 comparing the effects of different types of teacher evaluation on the well-being/ill-being  
4 of teachers. Fourth, the current study focused narrowly on how altered teacher  
5 motivation due to perceived evaluation pressures affects teachers' reported levels of  
6 vitality and exhaustion; however, there are many other dependent variables that could  
7 be explored, including objective records of teachers' health, teachers' turnover  
8 decisions, teachers' interpersonal behaviors and student motivation or student  
9 engagement. Finally, the results concerning amotivation should be viewed with caution,  
10 as the reliability index of this measure, the only significant mediator in this study, was  
11 marginally under the recommended cut-off value of .70.

12 Previous research (Berliner, 2014; Firestone, 2014; Weiner, 2004; Yuan et al.,  
13 2012) has suggested that decreases in teacher motivation and well-being could be  
14 related to the fact that this type of teacher evaluation focuses on metrics that are outside  
15 of the direct control of the teacher (e.g., the performance of their students). As such, an  
16 important direction for future research would be to explore whether teacher evaluation  
17 based on controllable (e.g., class preparation or teaching skills) or non-controllable  
18 aspects (e.g., student performance) differently affect motivation and well-being of  
19 teachers. It may be that more controllable types of evaluation can counteract the  
20 negative effects of less controllable ones by giving teachers opportunities to  
21 demonstrate their competencies. In addition, some types of evaluation only consider  
22 student performance in particular subjects (e.g., maths and reading). It would, therefore,  
23 be interesting to compare how this assessment affects teachers' psychological health  
24 depending on whether their subjects are included in the evaluation or not.

## 25 **Implications for Practice**

1           The results of the current study have demonstrated how teacher evaluation based  
2 on student performance can negatively affect teacher well-being, which has several  
3 potential implications for applied practice. One potential alternative to conducting  
4 teacher evaluation based on student performance would be to adopt a more holistic and  
5 inclusive definition of what is meant by good teaching. Good and Lavinge (2015), for  
6 example, emphasized that good teaching involves much more than increasing students'  
7 grades; for example, it means supporting students to become better problem solvers  
8 (promoting creativity and analytical skills) or stimulating students' civility and social  
9 responsibility (promoting respect and empathy). In other words, besides the traditional  
10 grades, teacher evaluations could also incorporate information about the degree of  
11 development of cognitive and social skills of students. As the results of the present  
12 study indicate that teacher evaluation based on student performance can negatively  
13 affect motivation and teacher well-being, school should consider the leadership styles of  
14 principals and administration managers with the emphasis being on helping them  
15 minimize coercive strategies, such as rewards and comparisons with others, and instead  
16 promoting autonomous teacher motivation for work (Eyal & Roth, 2011; Fernet et al.,  
17 2012).

18

19

20

21

22

**References**

- 1
- 2    Álvarez, M. S., Balaguer, I., Castillo, I., & Duda, J. L. (2012). The coach-created  
3       motivational climate, young athletes' well-being, and intentions to continue  
4       participation. *Journal of Clinical Sport Psychology*, *6*, 166-179. doi:  
5       [10.1123/jcsp.6.2.166](https://doi.org/10.1123/jcsp.6.2.166)
- 6    Balaguer, I., Castillo, I., Álvarez, M., & Duda, J. L. (2005). Importance of social  
7       context in the prediction of self-determination and well-being in athletes of  
8       different level. In *Symposium on motivation in sport and physical activity. CD-*  
9       *Rom of 9th European Congress of Psychology. Granada, Spain.*
- 10   Baker, E. L., Barton, P. E., Darling-Hammond, L., Haertel, E., Ladd, H. F., Linn, R.  
11     L., ... & Shepard, L. A. (2010). Problems with the use of student test scores to  
12     evaluate teachers. EPI Briefing Paper, 278. *Economic Policy Institute.*
- 13   Bartholomew, K. J., Ntoumanis, N., Cuevas, R., & Lonsdale, C. (2014). Job pressure  
14     and ill-health in physical education teachers: The mediating role of psychological  
15     need thwarting. *Teaching and Teacher Education*, *37*, 101-107. doi:  
16     [10.1016/j.tate.2013.10.006](https://doi.org/10.1016/j.tate.2013.10.006)
- 17   Berliner, D. C. (2014). Exogenous variables and value-added assessments: A fatal  
18     flaw. *Teachers College Record*, *116*, 1-31.
- 19   Betoret, F. D. (2006). Stressors, self-efficacy, coping resources, and burnout among  
20     secondary school teachers in Spain. *Educational Psychology*, *26*, 519-539. doi:  
21     [10.1080/01443410500342492](https://doi.org/10.1080/01443410500342492)
- 22   Blais, M. R., Brière, N. M., Lachance, L., Riddle, A. S., & Vallerand, R. J. (1993).  
23     L'inventaire des motivations au travail de Blais [Blais' Inventory of Motivations  
24     at Work]. *Revue québécoise de psychologie*, *14*(3), 185-215.

- 1 Blandford, S. (2000). *Managing professional development in schools*. London, UK:  
2 Routledge.
- 3 Bostic, T. J., Rubio, D. M., & Hood, M. (2000). A validation of the subjective  
4 vitality scale using structural equation modeling. *Social Indicators Research*, 52,  
5 313-324. doi: [10.1023/A:1007136110218](https://doi.org/10.1023/A:1007136110218)
- 6 Byrne, B. M. (2001). *Structural equation modeling with AMOS: Basic concepts,*  
7 *applications, and programming*. L Erlbaum Associates, Mahwah, NJ.
- 8 Cuevas, R., Sanchez-Oliva, D., Bartholomew, K. J., Ntoumanis, N., & Garcia-Calvo,  
9 T. (2015). Adaptation and validation of the Psychological Need Thwarting Scale  
10 in Spanish physical education teachers. *Spanish Journal of Psychology*, 18, 1–9.  
11 doi: [10.1017/sjp.2015.56](https://doi.org/10.1017/sjp.2015.56)
- 12 Cuevas, R., Sánchez-Oliva, D., Contreras, O., Moreno, H., & García-Calvo, T.  
13 (2014). Adaptación del inventario de motivación en el trabajo al contexto  
14 educativo español [Adaptation of the motivational inventory at work to the  
15 Spanish educational context]. In A. Romero, T. Ramiro-Sánchez, & M. P.  
16 Bermúdez (Eds.), *Libro de resúmenes de II Congreso Internacional de Ciencias*  
17 *de la Educación y del Desarrollo*. Granada: Asociación de Psicología Conductual.
- 18 Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in*  
19 *human behavior*. New York: Academic Press.
- 20 Dworkin, A. G., & Tobe, P. F. (2014). The effects of standards based school  
21 accountability on teacher burnout and trust relationships: a longitudinal analysis.  
22 In D. V. Maele, P. B. Forsyth, & M. V. Houtte, (Eds), *Trust and school life: the*  
23 *role of trust for learning, teaching, leading, and bridging* (pp. 121-143).  
24 Netherlands: Springer..

- 1 Eyal, O. & Roth, G. (2011). Principals' leadership and teachers' motivation: Self-  
2 determination theory analysis. *Journal of Educational Administration*, 49, 256-  
3 275. doi: [10.1108/09578231111129055](https://doi.org/10.1108/09578231111129055)
- 4 Feng, L., Figlio, D. N., & Sass, T. (2010). *School accountability and teacher*  
5 *mobility*. Cambridge, USA: National Bureau of Economic Research.
- 6 Fernet, C., Guay, F., Senécal, C., & Austin, S. (2012). Predicting intraindividual  
7 changes in teacher burnout: The role of perceived school environment and  
8 motivational factors. *Teaching and Teacher Education*, 28, 514-525. doi:  
9 [10.1016/j.tate.2011.11.013](https://doi.org/10.1016/j.tate.2011.11.013)
- 10 Finnigan, K. S., & Gross, B. (2007). Do accountability policy sanctions influence  
11 teacher motivation? Lessons from Chicago's low-performing schools. *American*  
12 *Educational Research Journal*, 44, 594-630. doi: [10.3102/0002831207306767](https://doi.org/10.3102/0002831207306767)
- 13 Firestone, W. A. (2014). Teacher evaluation policy and conflicting theories of  
14 motivation. *Educational Researcher*, 43, 100-107. doi:  
15 [10.3102/0013189X14521864](https://doi.org/10.3102/0013189X14521864)
- 16 Gil-Monte, P. R. (2002). Validez factorial de la adaptación al español del Maslach  
17 Burnout Inventory-General Survey [The factorial validity of the Maslach Burnout  
18 Inventory-General Survey]. *Salud pública de México*, 44, 33-40.
- 19 Good, T. L., & Lavigne, A. L. (2015). Issues of teacher performance stability are not  
20 new: limitations and possibilities. *Education Policy Analysis Archives*, 23, 1-16.  
21 doi: [10.14507/epaa.v23.1916](https://doi.org/10.14507/epaa.v23.1916)
- 22 Goldhaber, D., & Hannaway, J. (2004). Accountability with a kicker: Observations  
23 on the Florida A+ accountability plan. *Phi Delta Kappan*, 85, 598-605. doi:  
24 [10.1177/003172170408500807](https://doi.org/10.1177/003172170408500807)

- 1 Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data*  
2 *analysis*. New Jersey, USA. Maxwell Macmillan: International Editions.
- 3 Harvey, L.(2004). The power of accreditation: views of academics. *Journal of*  
4 *Higher Education Policy and Management, 26*, 207–223. doi:  
5 [10.1080/1360080042000218267](https://doi.org/10.1080/1360080042000218267)
- 6 Hewitt, K. K. (2015). Educator evaluation policy that incorporates EVAAS value-  
7 added measures: Undermined intentions and exacerbated inequities. *Education*  
8 *Policy Analysis Archives, 23*. 1-49. doi: [10.14507/epaa.v23.1968](https://doi.org/10.14507/epaa.v23.1968)
- 9 Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural equation modelling:  
10 Guidelines for determining model fit. *Electronic Journal of Business Research*  
11 *Methods, 6*, 53-60.
- 12 Hu, L., & Bentler, P. M. (1999). Cut-off criteria for fit indexes in covariance structure  
13 analysis: Conventional criteria versus new alternatives. *Structural Equation*  
14 *Modeling: A Multidisciplinary Journal, 6*, 1-55. doi: [10.1080/10705519909540118](https://doi.org/10.1080/10705519909540118)
- 15 Isore, M. (2009). Teacher evaluation: current practices in OECD countries and a  
16 literature review. OECD Education Working Papers, 23. *OECD Publishing (NJI)*.
- 17 Jiang, J. Y., Spote, S. E., & Luppescu, S. (2015). Teacher perspectives on evaluation  
18 reform Chicago’s REACH students. *Educational Researcher, 44*, 105-116. doi:  
19 [10.3102/0013189X15575517](https://doi.org/10.3102/0013189X15575517)
- 20 Klusmann, U., Kunter, M., Trautwein, U., Lüdtke, O., & Baumert, J. (2008).  
21 Teachers' occupational well-being and quality of instruction: The important role  
22 of self-regulatory patterns. *Journal of Educational Psychology, 100*, 702-715. doi:  
23 [10.1037/0022-0663.100.3.702](https://doi.org/10.1037/0022-0663.100.3.702)
- 24 Konstantopoulos, S. (2014). Teacher effects, value-added models, and  
25 accountability. *Teachers College Record, 116*, 1-21.

- 1 Marina, J. A., Pellicer, C., & Manso, J. (2015). *Libro blanco de la profesión docente*  
2 *y su entorno escolar [White paper of the teaching profession and its school*  
3 *environment]*. Madrid, Spain: Ministerio de Educación, Cultura y Deporte.
- 4 McCaffrey, D. F., Lockwood, J. R., Koretz, D. M., & Hamilton, L. S. (2003).  
5 *Evaluating value-added models for teacher accountability. Monograph.* RAND  
6 Corporation. PO Box 2138, Santa Monica, CA 90407-2138.
- 7 McCaffrey, D.F., Sass, T.R., Lockwood, J.R., & Kata, M. (2009). The intertemporal  
8 variability of teacher effect estimates. *Education Finance and Policy*, 4, 572-606.  
9 doi: [10.1162/edfp.2009.4.4.572](https://doi.org/10.1162/edfp.2009.4.4.572)
- 10 Ministry of Education, Culture, and Sport (2014). *TALIS 2013 International study of*  
11 *teaching and learning. Spanish Report.* Madrid: Subdirección General de  
12 Documentación y Publicaciones.
- 13 Organic Law for the Improvement of the Educative Quality 8/2013, 9<sup>th</sup> December  
14 (2013). [Ley Orgánica para la Mejora de la Calidad Educativa]. *Spanish State*  
15 *Official Bulletin, [Boletín oficial del Estado de España]*, 295.
- 16 Organisation for Economic Co-operation and Development (2014). *New insights*  
17 *from TALIS 2013. Teaching and learning in primary and upper secondary*  
18 *education.* Paris, France: OECD Publishing. doi:[10.1787/9789264226319-en](https://doi.org/10.1787/9789264226319-en)
- 19 Pas, E. T., Bradshaw, C. P., & Hershfeldt, P. A. (2012). Teacher-and school-level  
20 predictors of teacher efficacy and burnout: Identifying potential areas for  
21 support. *Journal of School Psychology*, 50, 129-145.  
22 doi:[10.1016/j.jsp.2011.07.003](https://doi.org/10.1016/j.jsp.2011.07.003)
- 23 Pastor, V. M. L., D. P., Arribas, J. C. M., & Aguado, R. M. (2016). Los retos de la  
24 Educación Física en el Siglo XXI [The challenges of physical education in the  
25 21st century]. *Retos*, 29, 182-187.

- 1 Pelletier, L. G., Séguin-Lévesque, C., & Legault, L. (2002). Pressure from above and  
2 pressure from below as determinants of teachers' motivation and teaching  
3 behaviors. *Journal of Educational Psychology*, *94*, 186-196. doi: [10.1037/0022-](https://doi.org/10.1037/0022-0663.94.1.186)  
4 [0663.94.1.186](https://doi.org/10.1037/0022-0663.94.1.186)
- 5 Preacher, K. J., & Hayes, A. F. (2008). Contemporary approaches to assessing  
6 mediation in communication research. In A. F. Hayes, M. D. Slater, y L. B.  
7 Snyder (Eds.), *The Sage sourcebook of advanced data analysis methods for*  
8 *communication research* (pp. 13-54). Thousand Oaks, CA: Sage Publications  
9 doi:[10.4135/9781452272054.n2](https://doi.org/10.4135/9781452272054.n2)
- 10 Quintana, S. M., & Maxwell, S. E. (1999). Implications of recent developments in  
11 structural equation modeling for counseling psychology. *The Counseling*  
12 *Psychologist*, *27*, 485–527. doi: [10.1177/0011000099274002](https://doi.org/10.1177/0011000099274002)
- 13 Rockoff, J. E. (2004). The impact of individual teachers on student achievement:  
14 Evidence from panel data. *The American Economic Review*, *94*, 247-252.
- 15 Roth, G., Assor, A., Kanat-Maymon, Y., & Kaplan, H. (2007). Autonomous motivation  
16 for teaching: How self-determined teaching may lead to self-determined learning.  
17 *Journal of Educational Psychology*, *99*, 761-774. doi: [10.1037/0022-0663.99.4.761](https://doi.org/10.1037/0022-0663.99.4.761)
- 18 Rothstein, J. (2010). Teacher quality in educational production: Tracking, decay, and  
19 student achievement. *Quarterly Journal of Economics*, *125*, 175-214. doi:  
20 [10.1162/qjec.2010.125.1.175](https://doi.org/10.1162/qjec.2010.125.1.175)
- 21 Ryan, R. M. & Frederick, C. (1997). On energy, personality, and health: Subjective  
22 vitality as a dynamic reflection of well-being. *Journal of Personality*, *65*, 529-  
23 565. doi: [10.1111/j.1467-6494.1997.tb00326.x](https://doi.org/10.1111/j.1467-6494.1997.tb00326.x)
- 24 Schaufeli, W. B., Leiter, M. P., Maslach, C., & Jackson, S. E. (1996). The MBI-  
25 general survey. In C. Maslach, S. E. Jackson, & M. P. Leiter (Eds.), *Maslach*

- 1 *Burnout inventory. Manual*, (pp. 19-26). Palo Alto, CA: Consulting Psychologists  
2 Press.
- 3 Tabachnick, B.G. & Fidell, L.S. (2007). *Using multivariate statistics* (5<sup>a</sup> ed.). New  
4 York: Allyn and Bacon.
- 5 Taylor, E. S., & Tyler, J. H. (2012). The effect of evaluation on teacher performance.  
6 *American Economic Review*, *102*, 3628-3651. doi: [10.1257/aer.102.7.3628](https://doi.org/10.1257/aer.102.7.3628)
- 7 Taylor, I., & Ntoumanis, N. (2007). Teacher motivational strategies and student self-  
8 determination in physical education. *Journal of Educational Psychology*, *99*, 747-  
9 760. doi: [10.1037/0022-0663.99.4.747](https://doi.org/10.1037/0022-0663.99.4.747)
- 10 Taylor, I., Ntoumanis, N., & Standage, M. (2008). A self-determination theory  
11 approach to understanding antecedents of teachers' motivational strategies in  
12 physical education. *Journal of Sport and Exercise Psychology*, *30*, 75-94. doi:  
13 [10.1123/jsep.30.1.75](https://doi.org/10.1123/jsep.30.1.75)
- 14 Vansteenkiste, M., Lens, W., De Witte, S., De Witte, H., & Deci, E. L. (2004). The  
15 'why' and 'why not' of job search behaviour: their relation to searching,  
16 unemployment experience, and well-being. *European Journal of*  
17 *Social Psychology*, *34*, 345-363. doi: [10.1002/ejsp.202](https://doi.org/10.1002/ejsp.202)
- 18 von der Embse , N. P., Pendergast, L. L., Segool, N., Saeki, E., & Ryan, S. (2016).  
19 The influence of test-based accountability policies on school climate and teacher  
20 stress across four states. *Teaching and Teacher Education*, *59*, 492-502. doi:  
21 [10.1016/j.tate.2016.07.013](https://doi.org/10.1016/j.tate.2016.07.013)
- 22 Weiner, B. (2004). Attribution theory revisited: Transforming cultural plurality into  
23 theoretical unity. In D. M. McInerney & S. Van Etten (Eds.), *Big theories*  
24 *revisited* (pp. 13–29). Greenwich, CT: Information Age Publishing.

- 1 Yuan, K., Le, V. N., McCaffrey, D. F., Marsh, J. A., Hamilton, L. S., Stecher, B. M.,
- 2 & Springer, M. G. (2012). Incentive pay programs do not affect teacher
- 3 motivation or reported practices results from three randomized studies.
- 4 *Educational Evaluation and Policy Analysis*, 35, 3-22.
- 5 doi:[10.3102/0162373712462625](https://doi.org/10.3102/0162373712462625)

1 Table 1

2 *Descriptive Statistics, Reliability Estimates and Pearson Correlations*

Variable	Range	<i>M</i>	<i>SD</i>	$\alpha$	CR	AVE	1	2	3	4	5
1. Perceived pressure	1 - 7	2.78	0.92	0.72	0.70	0.52					
2. Autonomous motivation	1 - 7	5.31	0.96	0.87	0.82	0.62	0.03				
3. Controlled motivation	1 - 7	4.24	1.05	0.75	0.73	0.53	0.15**	0.48**			
4. Amotivation	1 - 7	2.38	1.04	0.67	0.65	0.50	0.33**	-0.31**	0.05		
5. Vitality	1 - 7	5.55	1.22	0.94	0.93	0.69	-0.25**	0.41**	0.08	-0.44**	
6. Exhaustion	1 - 7	2.75	1.45	0.90	0.89	0.65	0.39**	-0.30**	0.02	0.53**	-0.61**

3 \*\*  $p < 0.01$ .4 *Note.*  $\alpha$  = Cronbach's alpha; CR = Reliability composite; AVE = Average variance extracted

5

6

7

8

1 Table 2

2 *Standardized Direct and Indirect Effects*

Independent variable	Criterion variable	Total direct effect (95% CI)	Total indirect effect (95% CI)	Specific indirect effects		
				Autonomous motivation (95% CI)	Controlled motivation (95% CI)	Amotivation (95% CI)
Perceived pressure	Vitality	-0.22* (0.34 to 0.97)	-0.10* (0.20 to 0.02)	0.02 (-0.04 to 0.07)	-0.01 (-0.03 to 0.01)	-0.11* (-0.18 to -0.06)
Perceived Pressure	Exhaustion	0.33* (0.19 to 0.47)	0.19* (0.10 to 0.31)	-0.01 (-0.06 to 0.03)	0.01 (-0.02 to 0.04)	0.20* (0.13 to 0.30)

3 *Note.* CI= 95% Confidence Intervals, \*= CI does not include zero

4

5

6

7

8

1  
2

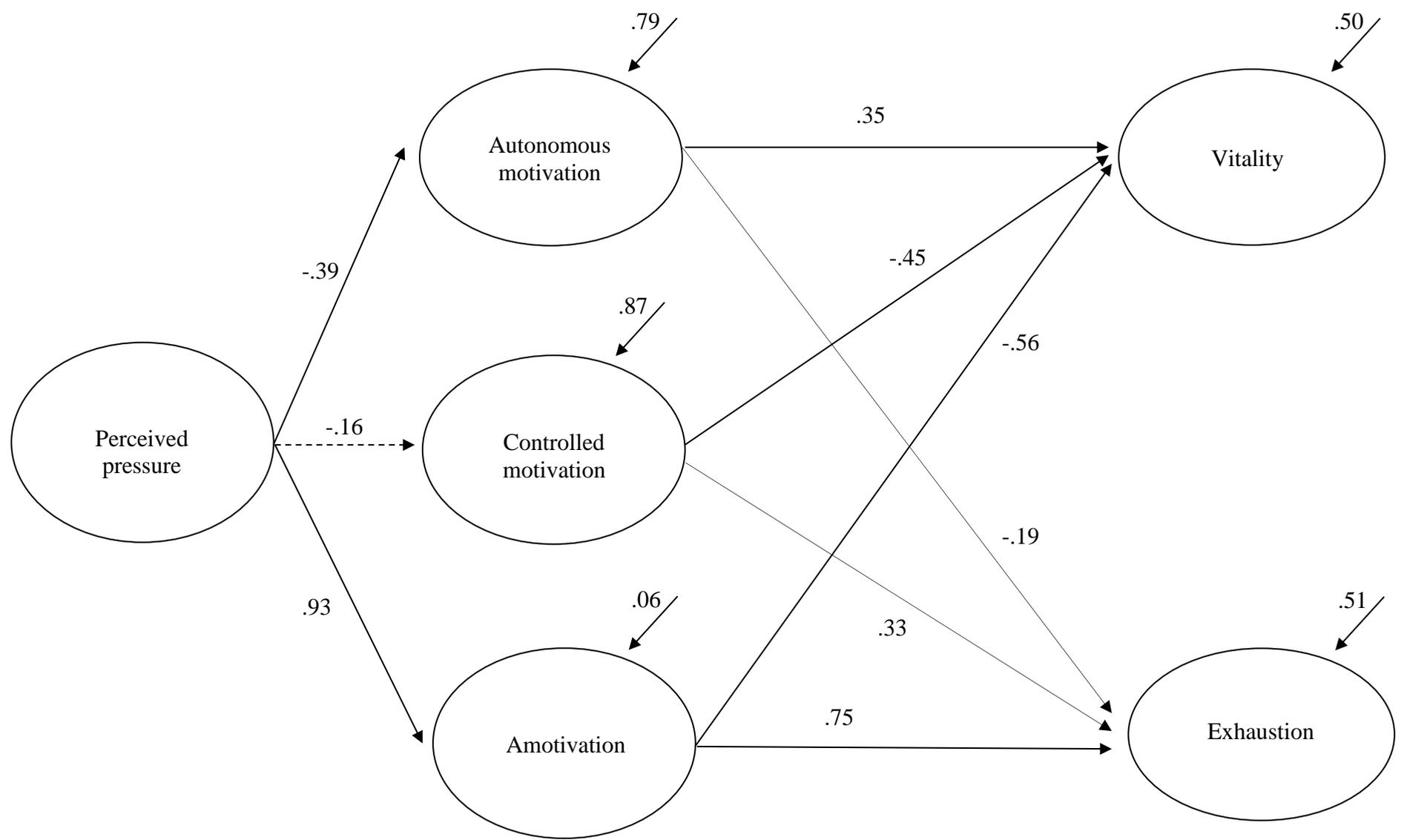


Fig. 1. Final model for the prediction of teacher well-being and teacher ill-being.  
Note: Significant paths are marked by solid lines. The small arrows over the dependent variables represent residual variance.