From heartbeat to data – Using wearable fitness trackers as an affordable approach to assess teacher stress

Mandy Klatt[a,1,*], Peer Keßler[a,2], Gregor Kachel[a,1,1], Christin Lotz[a,1], Anne Deiglmayr[a,3]

[a] *Empirische Schul- und Unterrichtsforschung, Marschnerstr. 29, Leipzig, 4109, Sachsen, Germany*
[b] *Department, Deutscher Pl. 6, Leipzig, 4103, Sachsen, Germany*

Abstract

Past research on physiological indicators of teacher stress often had to rely on expensive and obtrusive assessment methods. Modern fitness trackers represent a non-invasive and convenient alternative. This study investigated the use of wrist-worn fitness trackers to assess teacher heart rate (HR) as an indicator of stress during teaching. In a laboratory study, we used a Fitbit® fitness tracker to assess teachers´ HR before, during, and after a potentially stressful micro-teaching session. Our results demonstrate that the fitness tracker was indeed useful for mapping teachers' stress, with the data showing how teachers' HR increased before, peaked during, and progressively decreased after the micro-teaching session. Moreover, we related the fitness tracker data to retrospective teacher self-reports. We found that teachers' subjective stress appraisals, together with their teaching experience, explained only small amounts of variance in HR data. We discuss the potential of fitness trackers as an affordable and ubiquitous assessment tool for research on teacher stress in the classroom and provide advice for practical implementation.

Keywords: teacher stress, fitness tracker, heart rate, classroom disruptions, wearable technology, physiological stress measurement

1. Introduction

The teaching profession is one of the most stressful professions, with teachers facing a host of stressors during their everyday work (Smith, 2000; Herman et al., 2020; Schult et al., 2014). To better understand mechanisms in teacher stress, there is a growing research interest in physiological measures such as heart rate (HR) as online measures of teachers' stress during teaching (Kärner and Höning, 2021; Wettstein et al., 2020). For example, it has been shown that teacher-centered activities and typical classroom-related stressors increase teacher HR during teaching activities (Sperka and Kittler, 1995; Scheuch and Knothe, 1997; Donker et al., 2018; Junker et al., 2021; Huang et al., 2022). However, previous studies have often relied on expensive and obtrusive electrocardiographs (ECG). Modern fitness trackers represent a non-invasive and convenient alternative. (Ferguson et al., 2015)

Classroom disruptions are a major stressor in teachers' daily work (Boyle et al., 1995; Aloe et al., 2014), and learning how to deal with them is an important aspect of professional expertise (Wolff et al., 2015). According to Lazarus (1990) transactional model of stress and coping, the experience of stress in response to stressors such as classroom disruptions depends on the teacher´s subjective appraisal, which, in turn, depends on their coping resources, such as their professional knowledge. The resulting stress response has a

* Corresponding author
Email addresses: mandy.klatt@uni-leipzig.de (Mandy Klatt), Peer@example.com (Peer Keßler), gregor@example.com (Gregor Kachel), christin@example.com (Christin Lotz), anne@example.com (Anne Deiglmayr)
[1] ConceptualizationWriting - Original Draft PreparationWriting - Review & Editing
[2] Writing - Original Draft PreparationWriting - Review & Editing
[3] Supervision

psychological, physiological, or behavioral dimension (Kyriacou and Sutcliffe, 1978). Therefore, in order to better understand how classroom stressors affect teachers' stress response, subjective self-reports should be accompanied by objective, physiological measures (Wettstein et al., 2021). Teachers' use of wrist-worn fitness trackers in educational research provides fine-grained, in vivo data, allowing researchers as well as teachers themselves to monitor their physiological stress response continuously during teaching, across settings, and at low costs. Being able to monitor, and eventually counteract, teacher stress levels appears particularly relevant given the profession's generally high stress levels and associated negative health effects (Johnson et al., 2005; Montgomery and Rupp, 2005). To harness this potential, the present study explored the use of wrist-based fitness trackers as a tool to assess teachers' HR, as an indicator of stress, before, during, and after a teaching session during which typical, potentially stressful, classroom disruptions occurred. Further, we explored to what extent teachers' subjective appraisals of classroom disruptions and their teaching experience predicted teacher stress as assessed by the fitness tracker.

1.1. Fitness trackers as a ubiquitous, low-cost tool for assessing physiological stress responses

Fitness trackers provide data on physical activity and cardiovascular parameters such as HR, supporting personalized fitness goals (Nuss et al., 2021) and stress management (Hao et al., 2018). They can be used as ubiquitous, low-cost, and unintrusive data collection instruments (Godfrey et al., 2018), and their widespread use and everyday availability align with the increasing popularity and acceptance of wearables among the general population (Peng et al., 2022). In contrast to self-reported questionnaires on stress (Chaplain, 2008; Liu and Yan, 2020) that are prone to biases like social desirability (Razavi, 2001) or recall errors (Van den Bergh and Walentynowicz, 2016), fitness trackers, as ambulatory assessment methods (Trull and Ebner-Priemer, 2013; Wettstein et al., 2020), offer more objective insights into teachers' stress levels by monitoring teachers' physiological stress responses without disrupting teaching (Donker et al., 2018; Runge et al., 2020).

One important health parameter assessed by nearly all wrist-worn fitness trackers is heart rate (Scalise and Cosoli, 2018). HR indicates the number of heartbeats within one minute and is typically expressed as beats per minute (BPM) (Berntson et al., 2007; Hottenrot, 2007). At rest, the average HR of adults typically ranges between 60 and 80 BPM (Sammito et al., 2015). HR can be detected and measured in different ways using sensors, such as electrocardiography (ECG) or photoplethysmography (PPG) (Mukhopadhyay and Islam, 2017). While ECG sensors offer precise measurements by detecting the heart's electrical activity, their intrusive nature and requirement of direct skin contact may limit their suitability (Kranjec et al., 2014), particularly in educational settings. In contrast, photoplethysmography (PPG) is a rather uncomplicated and inexpensive technique to measure HR, commonly found in commercially available fitness trackers (Castaneda et al., 2018). This optical method assesses HR by flashing green or red lights to measure changes in blood volume in the capillaries of the skin (Allen, 2007).

Physiologically, HR is regulated by the sympathetic and parasympathetic nervous systems (Pham et al., 2021). An increase in sympathetic activity results in HR being sped up ("fight or flight" response) (Taelman et al., 2009); whereas an increase in parasympathetic activity results in HR being slowed down ("rest and digest" response) (Battipaglia and Lanza, 2015). Stress or mental and physical strain directly increases HR (Custodis et al., 2014; Sachs, 2014). Therefore, an increase in HR can be regarded as an indicator of increasing stress, and a decrease as an indicator of relaxation and ease (Kyriacou and Sutcliffe, 1978). Thus, fitness trackers offer low-cost and unobtrusive access to psychological stress data.

1.2. HR in teaching-learning contexts

Prior research, typically using ECG methods, has shown that changes in teachers' HR can be mapped onto stressors they experience during teaching. For example, teachers' HR tends to increase when teachers take an exposed position in student-teacher interaction (Sperka and Kittler, 1995; Scheuch and Knothe, 1997; Donker et al., 2018; Junker et al., 2021). Sperka and Kittler (1995) for example recorded the HR of 16 pre-service teachers during their first lesson and showed that teachers' HR increased significantly during teaching. The activation was particularly prominent at the beginning of the lesson and decreased over the course of the lesson. The authors suggested that pre-service teachers' proactive coping strategies, such as actively managing student interactions, helped lower their HR levels. Other ECG studies identified typical stressors predicting increases in HR, such as class size (Huang et al., 2022), or low student engagement and motivation (Junker et al., 2021). Junker et al. (2021) recorded the HR of 40 teachers during a real classroom lesson. Again, teacher stress, induced by factors such as low student engagement (e.g., lack of motivation or interest in tasks) or teacher-centered activities (e.g., teacher-focused classroom activities), resulted in elevated HR.

More recent studies have begun using wrist-worn fitness trackers to investigate HR trends in instructional settings (Darnell and Krieg, 2019; Chalmers et al., 2021). Darnell and Krieg (2019) measured the HR of 15 medical college students listening to lectures, using wrist-worn devices. The analysis revealed a constant decrease in HR throughout the lecture, with HR peaks during more interactive learning phases. Chalmers et al. (2021) used HR data from a fitness tracker to identify physiological changes during stress-inducing tasks (i.e., the Trier Social Stress Test; Kirschbaum et al. (1993)]. Average HR increased significantly from the resting to the stress inducing phases. Even though the participants of these previous studies (Darnell and Krieg, 2019; Chalmers et al., 2021) were not teachers but learners, the results are relevant for studying teacher stress as they demonstrated that HR can be effectively recorded using fitness trackers over the course of a whole learning unit, and HR changes are in line with the occurrence of activating or stress-inducing tasks.

To the best of our knowledge, only one study has directly assessed teachers' HR during teaching using a wrist-worn fitness tracker: Runge et al. (2020) assessed HR as an indicator of stress in four in-service teachers during authentic lessons. They used fitness trackers' recordings to create a profile for each teacher, with the aim of differentiating between teachers reporting higher vs. lower levels of stress. It was found that the combination of a high HR, a high number of steps, and short sleep duration was characteristic for teachers reporting high stress levels. However, it should be noted that the generalizability of these results is limited due to the small sample size.

In summary, previous studies have revealed that teachers' (and students') HR changes depend on their activities and the stressors they experience, with an increase in HR already before the expected stressors occur, and with peaks in activating phases (Darnell and Krieg, 2019; Chalmers et al., 2021). For teachers, teacher-centered phases led to an increase in HR (Sperka and Kittler, 1995; Scheuch and Knothe, 1997; Donker et al., 2018; Junker et al., 2021). However, there is a lack of studies using teacher-worn fitness trackers in larger samples, exploring the feasibility of this convenient tool for researching links between teachers' HR and stress inducing events and mechanisms.

1.3. A model of teacher stress

According to Kyriacou and Sutcliffe (1978), teacher stress can be defined as a negative affective response, typically accompanied by physiological changes such as increased HR, triggered by job-related demands, and perceived as threatening to one's self-esteem or well-being. Coping mechanisms help to reduce the perceived threat. Kyriacou's definition of teacher stress (see Kyriacou and Sutcliffe (1978) and, for a more recent adaptation, Van Dick (2006)) is based on the transactional stress model by Lazarus and colleagues [lazarus1966psychological; Lazarus (1990)], which highlights the interaction between an individual and the environment, whereby stress refers to a person's subjective reaction to an event (a stressor) that exceeds the person's adaptive resources.

Fig. 1 shows, in a simplified way, how classroom events affect teachers' stress level, according to the adaptation of the Lazarus model to teacher stress proposed by Van Dick (2006): When potential stressors (e.g., classroom disruptions) occur during teaching (1), teachers intuitively judge how disruptive the event is in a primary appraisal (2). If potential stressors are judged as threatening, i.e., as actual stressors (3), teachers consider whether they have sufficient resources for coping with the stressors (4). Teachers utilize these resources in trying to cope with the stressors, e.g., by employing classroom management strategies (5). In cases when coping fails, stress ensues, often accompanied by physiological reactions like increased HR (6). As part of the coping process, and dependent on its outcomes, teachers re-evaluate the stressor (7).

As shown in Fig. 1, teachers' primary and secondary appraisal as well as coping attempts are influenced by professional experience. As professional experience grows, teachers develop cognitive scripts for managing

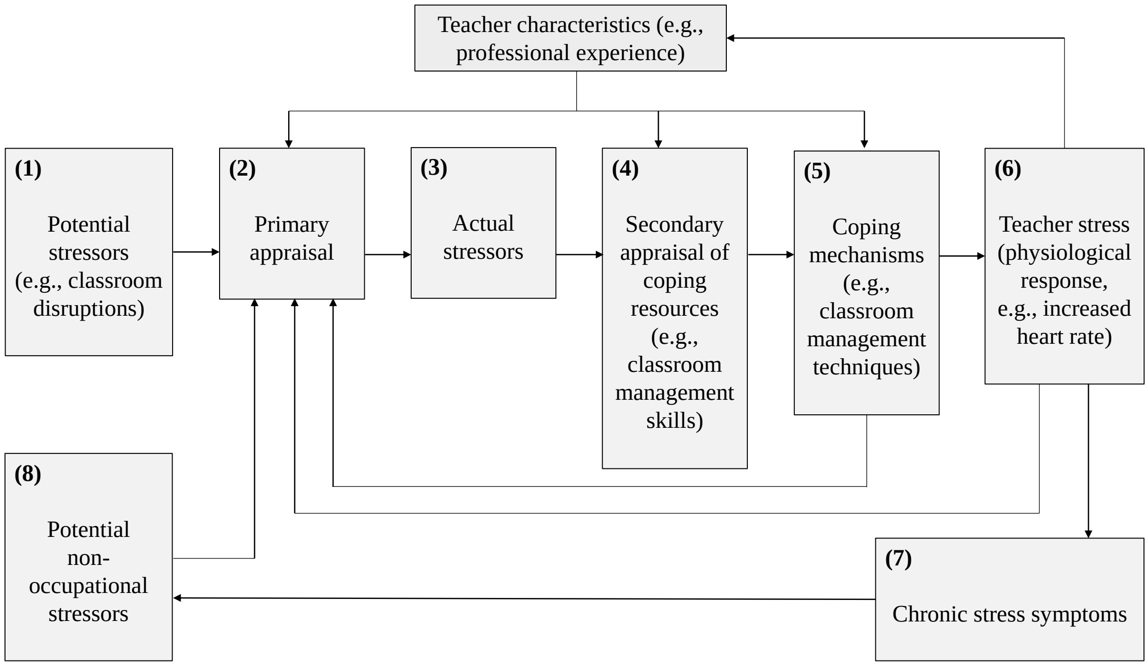


Figure 1: A model of teacher stress (adapted from van Dick 2006, p.37, modified by the authors)

classroom events, resulting in more complex and problem-focused classroom management skills (Wolff et al., 2021), and thus more effective coping and less stress. Empirically, classroom management skills and problem-focused coping styles are linked to fewer instances of emotional exhaustion (Maslach et al., 2001; Clunies-Ross et al., 2008). Novices in the teaching profession, on the other hand, face considerable stress and often feel overwhelmed by the demands of teaching (Ophardt and Thiel, 2017; Wolff et al., 2015; Klusmann et al., 2012) with many leaving the profession within the first five years (Ingersoll and Smith, 2003). Accordingly, when resources are lacking and coping fails, negative consequences for health (e.g., burnout) and for work (e.g., high turnover rates) can arise (Jalongo and Heider, 2006; Unterbrink et al., 2007; Aloe et al., 2014), highlighting the importance of professional expertise for managing teacher stress (Fisher, 2011).

1.4. Present Study

The present study aimed to explore the relations between teachers' HR response, and their subjective appraisals of stress during a micro-teaching unit, and to relate their self-reported appraisals and physiological stress responses to their teaching experience. We analyzed data from in-service and pre-service teachers who participated in a laboratory study as part of a larger project targeting the development of classroom management skills. Participants came to the lab individually and taught a short lesson to a class of three actors (i.e., trained student assistants) who performed several typical and possibly disruptive classroom events. The micro-teaching unit was thus potentially stressful for the participants. The aims of the present study were twofold:

(1) The first research goal was to investigate whether HR measures assessed by a wrist-based fitness tracker were a suitable and effective method for mapping teachers' HR over the course of the lab study, with a total duration of approximately 2 hours, including phases before, during, and after the stressful micro-teaching unit.

Looking at HR measures globally, we expected the participants to show an initial increase in their HR, followed by a peak during the micro-teaching unit and a decrease for the remaining phases. In addition, we examined whether z-standardization of the participants' HR could serve as a useful method to account for individual differences in baseline HR: We expected to observe the same trends in both standardized and non-standardized HR values.

In addition, we selected five representative 10-minute intervals from the five phases of the lab study (see Figure 2) in order to test the hypotheses that, regarding HR levels, teachers' HR would be the highest during the micro-teaching unit, compared to all other phases (**Hypothesis 1a**), and, regarding HR slopes, that teachers' HR would increase while they were preparing for teaching (pre-teaching interval), but decrease in all of the following intervals, i.e. when they were habituating to and recovering from the stressful micro-teaching unit (**Hypothesis 1b**).

(2) We further explored whether teaching experience made a difference in how teachers' HR reacted to the classroom disruptions. We expected more experienced teachers to be less stressed by the classroom events (**Hypothesis 2a**). In addition, we examined the relations between teachers' subjective appraisals of the classroom events (specifically, the disruptiveness of the events, and their confidence in dealing with them) and teachers' HR level, beyond the explanatory power of teaching experience. We expected higher HR levels for teachers who felt more disrupted, regardless of their teaching experience (**Hypotheses 2b**), and lower HR levels for teachers who felt more confident, regardless of teaching experience (**Hypothesis 2c**). We hypothesized that each of the three predictors (*teaching experience, disruption appraisal, confidence appraisal*) uniquely contributed to explaining variance in teachers' HR levels (**Hypothesis 2d**). In addition, we exploratively ran analogous analyses for the *changes* in HR (i.e., slopes).

2. Method

2.1. Participants

The sample consisted of $N = 84$ pre- and in-service teachers from Germany, who were recruited via personal contacts, email lists, and flyers. The data of three participants was lost due to failed data transmission, yielding an analysis sample of $n_{total} = 81$ ($n_{total} = 52$ women, $n_{total} = 29$ men), including 40 pre-service and 41 in-service teachers. Participants had a mean age of 30.95 years ($SD = 10.90$; range: 19-60) and an average teaching experience of 5.64 years ($SD = 9.46$; range: 0-37).

2.2. Setting and Procedure

The study was carried out following the ethical standards and the approval of the University's Institutional Review Board. All participants were informed in detail about the aims of the study prior to testing. Participation was voluntary, not incentivized, and only took place after written consent had been given.

Each participant came to the lab for a period of approximately two hours in total, and each participant underwent the same phases (see Fig. 2): In the *pre-teaching phase*, the experimenter welcomed the participants and helped them put on the fitness tracker. This was followed by a warm-up session to familiarize the participants with the laboratory setting and the class. This phase took about 10-15 minutes and participants spent this time mostly standing or slowly walking around. During the *teaching phase*, the participants held their self-prepared micro-teaching unit to a class of three trained actors who performed nine, potentially disruptive, classroom events (e.g., chatting with a neighbor, heckling, looking at the phone; see Table ## in the supplementary material for an overview and categorization of all events; and Fig## in the supplementary material for a depiction of the laboratory setting of the micro-teaching unit). The topic and class level of the teaching unit could be freely chosen by the teachers with the only requirement that the unit had to be an introductory lesson, and had to consist of supervised individual work and / or frontal teaching. The micro-teaching unit lasted about 15-20 minutes. Participants spent this time mostly standing or slowly walking around. While teaching, participants wore eye-tracking glasses, and their lesson was video-recorded. After having completed the micro-teaching unit, in the *post-teaching phase*, participants filled in questionnaires for approximately 10-15 minutes: a brief computer-based survey of sociodemographic data (e.g., teaching experience, gender, studied school type, studied school subjects, extracurricular teaching activities), and a short knowledge test that was irrelevant to the present study. In the *interview phase*, participants engaged in a Stimulated Recall Interview (SRI). During the SRI, participants sat in front of a computer monitor and watched the video of their own lesson from the ego perspective, as recorded through the eye-tracking glasses. The experimenter stopped the video each time one of the nine classroom events happened, and asked five open-ended, and three rating questions per event. Two of the rating questions are relevant to the present study: the disruption and the confidence appraisal ratings (see Measures). The interview lasted about 45-60 minutes. Finally, in the *end phase*, participants filled in another questionnaire irrelevant to the present study, which lasted about 10-15 minutes.

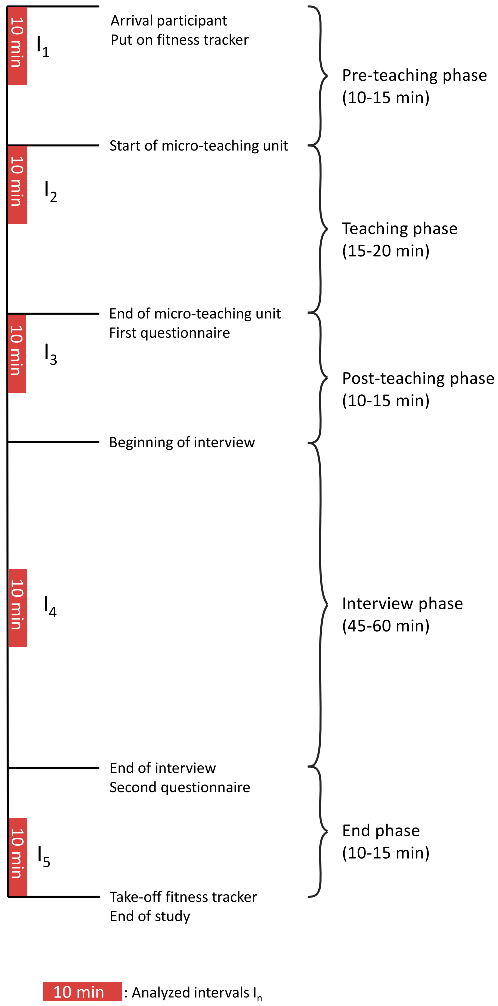


Figure 2: Procedure of the two-hour-long study consisting of five phases with five representative 10-minute intervals

2.3. Measures

2.3.1. Heart Rate Data and Heart Rate Intervals

To measure teachers' HR, we used the wrist-based fitness tracker Fitbit® Charge 4. In line with the manufacturer's instructions (, n.d.), the device was attached to the participants' non-dominant hand, a finger's width above the wrist bone. The tracker works by flashing green LEDs hundreds of times per second, using light-sensitive photodiodes to catch the reflected light, to calculate the volume changes in the capillaries. From this, the tracker calculated the heart beats per minute. HR measurements were generated at least every 15 seconds . The raw data contained the estimated HR in BPM for each time stamp. To account for individual differences in the baseline HR, we also calculated z-standardized HR values based on individual means, i.e., at the subject level of $n = 81$ participants (standardized HR).

Since we aimed to keep measurement intervals comparable between study phases, we aggregated HR over a representative 10-minute interval within each phase (cf. Fig. 2). Previous research has indicated that 10-minute intervals are a useful duration for analyzing PPG data (Lu et al., 2008). The intervals were selected based on the following rules: The *pre-teaching interval* (I_1) comprised the first 10 minutes after the fitness tracker had been put on. The *teaching interval* (I_2) started two minutes after the lesson had started. This interval was of the highest relevance to our study. We explicitly chose an early 10-minute interval within the teaching phase, as previous studies revealed that the beginning of a lesson is most demanding and potentially stressful with regards to teacher-student interaction (Donker et al., 2018; Claessens et al., 2017). The *post-teaching interval* (I_3) started immediately after the end of the teaching unit. The *interview interval* (I_4) was defined as the mid-10 minutes between the end of the teaching unit and the time point when the fitness tracker was taken off. All participants were being interviewed during this interval. The end interval (I5) comprised the last 10 minutes before the fitness tracker was taken off.

Interval	*M HR*	*SD HR*	Min	Max
Overall Course of 2h	90.09/0.04[4]	15.76/0.991	51/-4.03	164/4.56
Pre-teaching interval (I_1)	96.28/0.48	14.11/0.88	56/-3.56	139/3.24
Teaching interval (I_2)	100.80/0.85	16.23/0.77	63/-2.18	164/4.37
Post-teaching interval (I_3)	93.61/0.27	14.01/0.76	60/-2.17	150/3.06
Interview interval (I_4)	82.32/-0.72	11.85/0.74	51/-2.51	132/4.39
End interval I_5)	77.95/-1.07	11.14/0.57	50[5]/-2.68	120/2.96

Table 1: Mean HR (M), standard deviations HR (SD), and range of teachers' HR over the course of the entire study and the five intervals (unstandardized in BPM/z-standardized).

[-2] Please note that standardized M and SD of the overall course were not exactly 0 and 1 due to rounding differences

[-1] Deviations of the minimum values in the overall course vs. the end interval (I_5) are due to data of a few participants who needed more than two hours to finish the study.

[0] Please note that standardized M and SD of the overall course were not exactly 0 and 1 due to rounding differences

[1] Deviations of the minimum values in the overall course vs. the end interval (I_5) are due to data of a few participants who needed more than two hours to finish the study.

[2] Please note that standardized M and SD of the overall course were not exactly 0 and 1 due to rounding differences

[3] Deviations of the minimum values in the overall course vs. the end interval (I_5) are due to data of a few participants who needed more than two hours to finish the study.

[4] Please note that standardized M and SD of the overall course were not exactly 0 and 1 due to rounding differences

[5] Deviations of the minimum values in the overall course vs. the end interval (I_5) are due to data of a few participants who needed more than two hours to finish the study.

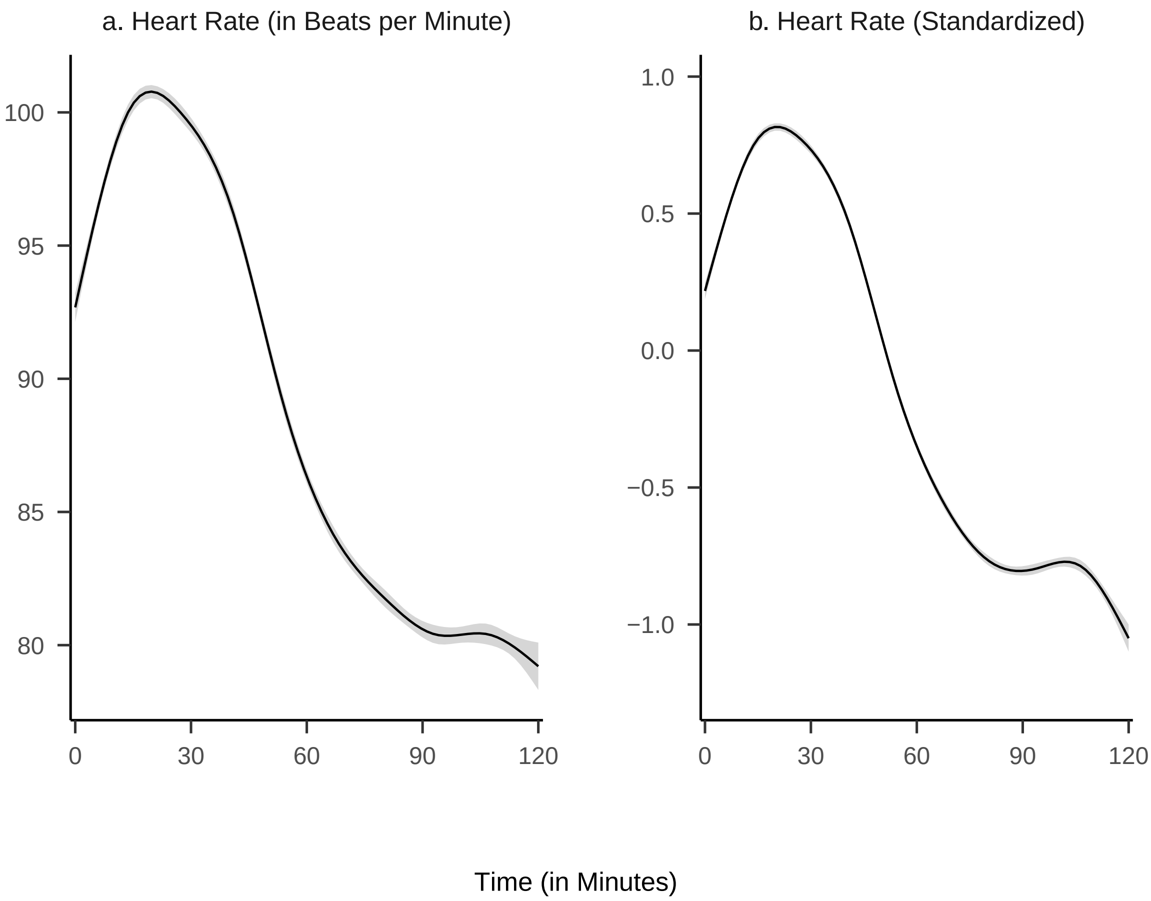
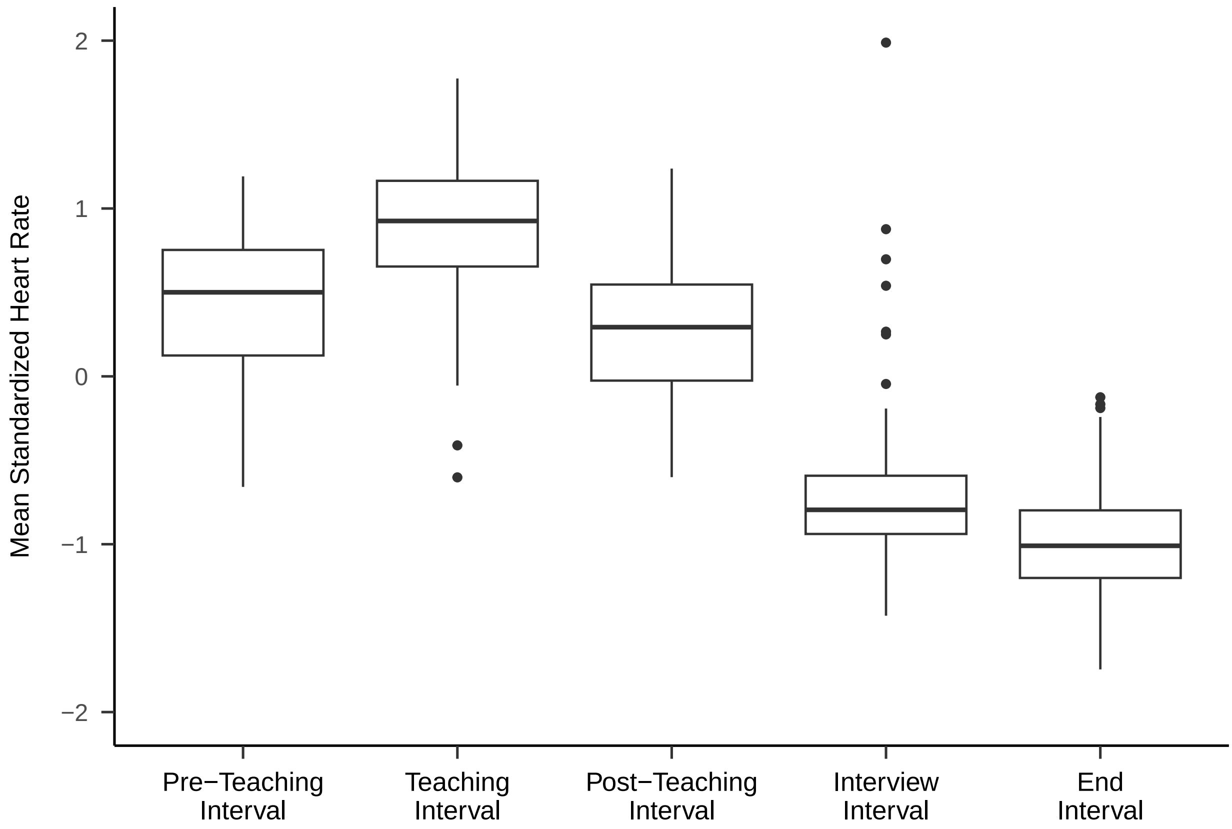


Figure 3: Overall course of the HR with the unstandardized HR in BPM shown in Fig. 3a. and the z-standardized HR shown in Fig. 3b. for the planned 2-hour study.



Note: N = 81 participants per interval. Fig. shows median (bold line), interquartile range (box) and outliers (dots).

Figure 4: Distribution of the standardized heart rate means in the five intervals

[4] All measurement points per interval for all participants. Note that the variation in n stems from the variation in the number of collected data points by the fitness tracker.

[5] All measurement points per interval for all participants. Note that the variation in n stems from the variation in the number of collected data points by the fitness tracker.

[6] All measurement points per interval for all participants. Note that the variation in n stems from the variation in the number of collected data points by the fitness tracker.

Interval	n[6]	M (SD)		p	
		Intercept	Slope	Intercept	Slope
(1) Pre-teaching interval	6896	0.052 (0.820)	0.085* (0.133)	.57	$<$ .05
(2) Teaching interval	7150	1.025* (0.690)	-0.039* (0.108)	$<$ .05	$<$ .05
(3) Post-teaching interval	6664	0.549* (0.547)	-0.060* (0.101)	$<$ .05	$<$ .05
(4) Interview interval	6287	-0.617* (0.614)	-0.022 (0.070)	$<$ .05	.006
(5) End interval	5990	-1.004* (0.500)	-0.012 (0.074)	$<$ .05	.14

Note. * $p < .05$

Table 2: Descriptive statistics (n, M, SD) for the mean intercepts and the mean slopes for the five intervals.

Variable	Pre-teaching interval	Teaching interval	Post-teaching interval	Interview interval	End interval
Teaching Experience	$-.17/-.27^{*}$	.11/-.02	$-.04/-.03$	$.24^{*}/-.20$	.04/.11
Disruption Appraisal	$-.01/.16$	$-.20/.08$	$.20/-.14$	$-.13/.01$	.04/.12
Confidence Appraisal	$-.10/-.18$	.06/.09	$.04/-.03$	$.09/-.19$	$-.07/.13$

Note. * $p < .05$.

Table 3: Correlations between mean standardized HR/mean slopes and the predictor variables of teaching experience (TE), disruption appraisal (DA), and confidence appraisal (CA) for the five intervals.

Dependent variable:	Model 1				Model 2				Model 3				Model 4			
	Mean standardized HR and mean slopes															
	Mean std. HR		Mean slopes		Mean std. HR		Mean slopes		Mean std. HR		Mean slopes		Mean std. HR		Mean slopes	
	β (SE)	p	β (SE)	p	β (SE)	p	β (SE)	p	β (SE)	p	β (SE)	p	β (SE)	p	β (SE)	p
Pre-teaching interval (I1)																
Teaching Experience	$-.17$ (.005)	.12	$-.27^*$ (.002)	$< .05$												
R^2	.030		.071													
Teaching interval (I2)																
Teaching Experience	.11 (.002)	.34	$-.02$ (.001)	.83	.04 (.005)	.73	.01 (.001)	.96	.10 (.006)	.42	$-.08$ (.001)	.54	.05 (.006)	.67	$-.05$ (.001)	.72
Disruption Appraisal	$-.18$ (.041)	.13	.08 (.010)	.50	$-.19$ (.042)	.13	.12 (.010)	.34								
Confidence Appraisal	.01 (.046)	.92	.12 (.011)	.34	$-.04$ (.047)	.76	.15 (.012)	.24								
R^2	.012		.000		.040		.015		.012		.010		.042		.031	
Δ R^2			.028		.015		.000		.010		.030		.031			

Continued on next page

Table 4 – continued from previous page

Dependent variable:	Model 1				Model 2				Model 3				Model 4			
	Mean standardized HR and mean slopes															
	β (SE)	p	β (SE)	p	β (SE)	p	β (SE)	p	β (SE)	p	β (SE)	p	β (SE)	p	β (SE)	p
Post-teaching interval (I3)																
Teaching Experience	$-.04$ (.005)	.70	$-.03$ (.001)	.80	.04 (.005)	.76	$-.09$ (.001)	.44	$-.08$ (.006)	.55	$-.02$ (.001)	.89	$-.01$ (.006)	.91	$-.07$ (.001)	.61
Disruption Appraisal	.22 (.040)	.07	$-.18$ (.009)	.14	.25* (.041)	$< .05$	$-.20$ (.010)	.12								
Confidence Appraisal	.08 (.045)	.55	$-.03$ (.011)	.83	.14 (.046)	.27	$-.08$ (.011)	.54								
R^2	.002		.001		.043		.020		.006		.002		.058		.023	
ΔR^2			.041		.019		.004		.001		.056		.022			
Interview interval (I4)																
Teaching Experience	.24* (.006)	$< .05$	$-.20$ (.001)	.07	.22 (.006)	.06	$-.23$ (.001)	.06	.25* (.006)	$< .05$	$-.14$ (.001)	.25	.23 (.007)	.07	$-.17$ (.001)	.18
Disruption Appraisal	$-.05$ (.045)	.66	$-.08$ (.006)	.52	$-.06$ (.047)	.61	$-.12$ (.007)	.34								

Continued on next page

Table 4 – continued from previous page

	Model 1				Model 2				Model 3				Model 4			
Dependent variable:	Mean standardized HR and mean slopes															
	β (SE)	p	β (SE)	p	β (SE)	p	β (SE)	p	β (SE)	p	β (SE)	p	β (SE)	p	β (SE)	p
Confidence Appraisal	−.02 (.050)	.85	−.13 (.007)	.29	−.04 (.052)	.76	−.16 (.007)	.20								
R^2	.058		.040		.060		.050		.058		.054		.061		.069	
ΔR^2			.002		.010		.000		.014		.003		.029			
End interval (I5)																
Teaching Experience	.04 (.004)	.70	.11 (.001)	.32	.07 (.005)	.58	.18 (.001)	.13	.09 (.005)	.46	.07 (.001)	.58	.10 (.005)	.43	.12 (.001)	.33
Disruption Appraisal	.06 (.035)	.60	.19 (.007)	.12	.04 (.037)	.76	.23 (.007)	.07								
Confidence Appraisal	−.11 (.039)	.38	.10 (.008)	.43	−.10 (.041)	.44	.16 (.008)	.22								
R^2	.002		.013		.005		.053		.012		.025		.013		.078	
ΔR^2			.003		.040		.010		.012		.011		.065			

Note. In Model 1, mean standardized HR and mean slopes were predicted only by teaching experience. In Model 2, solely disruption appraisal was added to teaching experience as a predictor. In Model 3, solely confidence appraisal was added to teaching experience as a predictor. In Model 4, all three predictors were considered in concert.
* $p < .05$.

References

. *Fitbit Charge 4 User Manual Version 1.2*. Fitbit, Incorp., n.d. URL `https://staticcs.fitbit.com/content/assets/help/manuals/manual_charge_4_en_US.pdf`.

John Allen. Photoplethysmography and its application in clinical physiological measurement. *Physiological measurement*, 28 (3):R1, 2007.

Ariel M Aloe, Shannon M Shisler, Benjamin D Norris, Amanda B Nickerson, and Tyler W Rinker. A multivariate meta-analysis of student misbehavior and teacher burnout. *Educational research review*, 12:30–44, 2014.

Irma Battipaglia and Gaetano A. Lanza. *The Autonomic Nervous System of the Heart*, pages 1–12. Springer Berlin Heidelberg, Berlin, Heidelberg, 2015. ISBN 978-3-662-45074-1. doi: 10.1007/978-3-662-45074-1_1. URL `https://doi.org/10.1007/978-3-662-45074-1_1`.

Gary G Berntson, Karen S Quigley, and D Lozano. Cardiovascular psychophysiology. *Handbook of psychophysiology*, 3:182–210, 2007.

Gregory J Boyle, Mark G Borg, Joseph M Falzon, and Anthony J Baglioni Jr. A structural model of the dimensions of teacher stress. *British Journal of Educational Psychology*, 65(1):49–67, 1995.

Denisse Castaneda, Aibhlin Esparza, Mohammad Ghamari, Cinna Soltanpur, and Homer Nazeran. A review on wearable photoplethysmography sensors and their potential future applications in health care. *International journal of biosensors & bioelectronics*, 4(4):195, 2018.

Taryn Chalmers, Blake Anthony Hickey, Phillip Newton, Chin-Teng Lin, David Sibbritt, Craig S McLachlan, Roderick Clifton-Bligh, John Morley, and Sara Lal. Stress watch: The use of heart rate and heart rate variability to detect stress: A pilot study using smart watch wearables. *Sensors*, 22(1):151, 2021.

Roland P Chaplain. Stress and psychological distress among trainee secondary teachers in england. *Educational Psychology*, 28(2):195–209, 2008.

Luce CA Claessens, Jan van Tartwijk, Anna C van der Want, Helena JM Pennings, Nico Verloop, Perry J den Brok, and Theo Wubbels. Positive teacher–student relationships go beyond the classroom, problematic ones stay inside. *The Journal of Educational Research*, 110(5):478–493, 2017.

Penny Clunies-Ross, Emma Little, and Mandy Kienhuis. Self-reported and actual use of proactive and reactive classroom management strategies and their relationship with teacher stress and student behaviour. *Educational psychology*, 28(6): 693–710, 2008.

F Custodis, JC Reil, SH Schirmer, O Adam, S Möhlenkamp, U Laufs, and M Boehm. Heart rate: Clinical variable and risk marker. *Deutsche Medizinische Wochenschrift (1946)*, 139(33):1661–8, 2014.

D. Darnell and P. Krieg. Student engagement, assessed using heart rate, shows no reset following active learning sessions in lectures. *PLoS ONE*, 14, 2019.

Monika H Donker, Tamara Van Gog, and M Tim Mainhard. A quantitative exploration of two teachers with contrasting emotions: intra-individual process analyses of physiology and interpersonal behavior. *Frontline Learning Research*, 6(3): 162–184, 2018.

Ty Ferguson, Alex V Rowlands, Tim Olds, and Carol Maher. The validity of consumer-level, activity monitors in healthy adults worn in free-living conditions: a cross-sectional study. *International journal of behavioral nutrition and physical activity*, 12 (1):1–9, 2015.

Molly H Fisher. Factors influencing stress, burnout, and retention of secondary teachers. *Current issues in education*, 14(1), 2011.

Alan Godfrey, Victoria Hetherington, Hubert Shum, Paolo Bonato, NH Lovell, and Sam Stuart. From a to z: Wearable technology explained. *Maturitas*, 113:40–47, 2018.

T Hao, H Chang, M Ball, K Lin, and X Zhu. chrv uncovering daily stress dynamics using bio-signal from consumer wearables. In *In MEDINFO 2017: Precision Healthcare Through Informatics: Proceedings of the 16th World Congress on Medical and Health Informatics*, volume 245, page 98, 2018.

Keith C Herman, Sara L Prewett, Colleen L Eddy, Alyson Savala, and Wendy M Reinke. Profiles of middle school teacher stress and coping: Concurrent and prospective correlates. *Journal of School Psychology*, 78:54–68, 2020.

K. Hottenrot. *Trainingskontrolle: mit Herzfrequenz-Messgeräten, 2. Auflage*. Meyer & Meyer Verlag, 2007.

Yizhen Huang, Eric Richter, Thilo Kleickmann, and Dirk Richter. Class size affects preservice teachers' physiological and psychological stress reactions: An experiment in a virtual reality classroom. *Computers & Education*, 184:104503, 2022.

Richard M Ingersoll and Thomas M Smith. The wrong solution to the teacher shortage. *Educational leadership*, 60(8):30–33, 2003.

Mary Renck Jalongo and Kelly Heider. Editorial teacher attrition: An issue of national concern. *Early Childhood Education Journal*, 33:379–380, 2006.

Sheena Johnson, Cary Cooper, Sue Cartwright, Ian Donald, Paul Taylor, and Clare Millet. The experience of work-related stress across occupations. *Journal of managerial psychology*, 20(2):178–187, 2005.

Robin Junker, Monika H Donker, and Tim Mainhard. Potential classroom stressors of teachers: An audiovisual and physiological approach. *Learning and Instruction*, 75:101495, 2021.

Tobias Kärner and Jana Höning. Teachers' experienced classroom demands and autonomic stress reactions: Results of a pilot study and implications for process-oriented research in vocational education and training. *Empirical Research in Vocational Education and Training*, 13:1–22, 2021.

Clemens Kirschbaum, Karl-Martin Pirke, and Dirk H Hellhammer. The 'trier social stress test'–a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology*, 28(1-2):76–81, 1993.

Uta Klusmann, Mareike Kunter, Thamar Voss, and Jürgen Baumert. Berufliche beanspruchung angehender lehrkräfte: Die effekte von persönlichkeit, pädagogischer vorerfahrung und professioneller kompetenz. *Zeitschrift für pädagogische Psychologie*, 2012.

Jure Kranjec, S Beguš, G Geršak, and J Drnovšek. Non-contact heart rate and heart rate variability measurements: A review. *Biomedical signal processing and control*, 13:102–112, 2014.

Chris Kyriacou and James Sutcliffe. Teacher stress: Prevalence, sources, and symptoms. *British journal of educational psychology*, 48(2):159–167, 1978.

Richard S Lazarus. Theory-based stress measurement. *Psychological inquiry*, 1(1):3–13, 1990.

Meihua Liu and Yi Yan. Anxiety and stress in in-service chinese university teachers of arts. *International journal of higher education*, 9(1):237–248, 2020.

Sheng Lu, He Zhao, Kihwan Ju, Kunson Shin, Myoungho Lee, Kirk Shelley, and Ki H Chon. Can photoplethysmography variability serve as an alternative approach to obtain heart rate variability information? *Journal of clinical monitoring and computing*, 22(1):23–29, 2008.

Christina Maslach, Wilmar B Schaufeli, and Michael P Leiter. Job burnout. *Annual review of psychology*, 52(1):397–422, 2001.

Cameron Montgomery and André A Rupp. A meta-analysis for exploring the diverse causes and effects of stress in teachers. *Canadian Journal of Education/Revue canadienne de l'éducation*, pages 458–486, 2005.

Subhas Chandra Mukhopadhyay and Tarikul Islam. *Wearable Sensors: Applications, design and implementation*. IOP Publishing, 2017.

Kayla Nuss, Kristen Moore, Tracy Nelson, and Kaigang Li. Effects of motivational interviewing and wearable fitness trackers on motivation and physical activity: A systematic review. *American Journal of Health Promotion*, 35(2):226–235, 2021.

Diemut Ophardt and Felicitas Thiel. Klassenmanagement als basisdimension der unterrichtsqualität. *Lehrer-Schüler-Interaktion: Inhaltsfelder, Forschungsperspektiven und methodische Zugänge*, pages 245–266, 2017.

Chenming Peng, Nannan Xi, Hong Zhao, and Juho Hamari. Acceptance of wearable technology: a meta-analysis. In *Hawaii International Conference on System Sciences*, pages 5101–5110, 2022.

Tam Pham, Zen Juen Lau, SH Annabel Chen, and Dominique Makowski. Heart rate variability in psychology: A review of hrv indices and an analysis tutorial. *Sensors*, 21(12):3998, 2021.

Tiffani Razavi. Self-report measures: An overview of concerns and limitations of questionnaire use in occupational stress research. 2001.

N Runge, S Haarman, and M Fisher. Using fitbit fitness trackers to measure teacher stress and coping. *International Journal of Social Policy and Education*, 2:56–70, 2020.

Stefanie Sachs. *Physiologische Parameter zur Bewertung der Lernwirksamkeit von Lernsituationen*. PhD thesis, Universität Ulm, 2014.

S Sammito, B Thielmann, R Seibt, A Klussmann, M Weippert, and I Böckelmann. Guideline for the application of heart rate and heart rate variability in occupational medicine and occupational science. *ASU Int*, 2015(06):1–29, 2015.

Lorenzo Scalise and Gloria Cosoli. Wearables for health and fitness: Measurement characteristics and accuracy. In *2018 IEEE international instrumentation and measurement technology conference (I2MTC)*, pages 1–6. IEEE, 2018.

Klaus Scheuch and Monika Knothe. Psychophysische beanspruchung von lehrern in der unterrichtstätigkeit. *Jahrbuch für Lehrerforschung*, 1(S 285):299, 1997.

Johannes Schult, Manuela Münzer-Schrobildgen, and Jörn R Sparfeldt. Belastet, aber hochzufrieden? *Zeitschrift für Gesundheitspsychologie*, 2014.

A Smith. The scale of perceived occupational stress. *Occupational Medicine*, 50(5):294–298, 2000.

M. Sperka and U.. Kittler. *Psychophysiologische Beanspruchung von Lehramtskandidatinnen und -kandidaten im Schulunterricht*, pages 182–197. Die Blaue Eule, Essen, 1995.

Joachim Taelman, Steven Vandeput, Arthur Spaepen, and Sabine Van Huffel. Influence of mental stress on heart rate and heart rate variability. In *4th European conference of the international federation for medical and biological engineering*, pages 1366–1369. Springer, 2009.

Timothy J Trull and Ulrich Ebner-Priemer. Ambulatory assessment. *Annual review of clinical psychology*, 9(1):151–176, 2013.

Thomas Unterbrink, Anna Hack, Ruth Pfeifer, Veronika Buhl-Grießhaber, Udo Müller, Helmut Wesche, Markus Frommhold, Klaus Scheuch, Reingard Seibt, Michael Wirsching, et al. Burnout and effort–reward-imbalance in a sample of 949 german teachers. *International archives of occupational and environmental health*, 80:433–441, 2007.

Omer Van den Bergh and Marta Walentynowicz. Accuracy and bias in retrospective symptom reporting. *Current opinion in psychiatry*, 29(5):302–308, 2016.

Rolf Van Dick. Stress und arbeitszufriedenheit bei lehrerinnen und lehrern. *Zwischen „Horrorjob "und Erfüllung*, 2, 2006.

Alexander Wettstein, Fabienne Kühne, Wolfgang Tschacher, and Roberto La Marca. Ambulatory assessment of psychological and physiological stress on workdays and free days among teachers. a preliminary study. *Frontiers in neuroscience*, 14: 495378, 2020.

Alexander Wettstein, Sandra Schneider, Martin grosse Holtforth, and Roberto La Marca. Teacher Stress: A Psychobiological Approach to Stressful Interactions in the Classroom. *Frontiers in Education*, 6, 2021. ISSN 2504-284X. URL `https://www.frontiersin.org/articles/10.3389/feduc.2021.681258`.

Charlotte E Wolff, Niek van den Bogert, Halszka Jarodzka, and Henny PA Boshuizen. Keeping an eye on learning: Differences between expert and novice teachers' representations of classroom management events. *Journal of teacher education*, 66(1): 68–85, 2015.

Charlotte E Wolff, Halszka Jarodzka, and Henny PA Boshuizen. Classroom management scripts: A theoretical model contrasting expert and novice teachers' knowledge and awareness of classroom events. *Educational Psychology Review*, 33(1): 131–148, 2021.