A Quantitative Exploration of Two Teachers with Contrasting Emotions: Intra-Individual Process Analyses of Physiology and Interpersonal Behavior

Monika H. Donker, Tamara van Gog, & M. Tim Mainhard

Department of Education, Utrecht University, The Netherlands

Abstract

Although the association between teacher-student relations, teacher emotions, and burnout has been proven on a general level, we do not know the exact processes underlying these associations. Recently there has been a call for intra-individual process measures that assess what happens from moment-to-moment in class in order to better understand inter-individual differences in emotions and burnout between teachers. This paper explored the use of process measures of teachers’ heart rate and their interpersonal behavior during teaching. Our aim was to illustrate different ways of analyzing and combining physiological and observational time-series data and to explore their potential for understanding between-teacher differences. In this illustration, we focused on two teachers who represented contrasting cases in terms of their self-reported teaching-related emotions (i.e., anxiety and relaxation) and burnout. We discuss both univariate process analyses (i.e., trend, autocorrelation, stability) as well as state-of-the-art multivariate process analyses (i.e., cross-correlations, dynamic structural equation modeling). Results illustrate how the two teachers differed in the nature of their physiological responses, their interpersonal behavior, and the association between these two process measures over time. Along implications and suggestions for further research, it is discussed how the process-based, dynamic assessment of physiology and interpersonal behavior may ultimately help to understand differences in more general teaching-related emotions and burnout.

Keywords: Teacher emotions; process analyses; physiology; heart rate; interpersonal teacher behavior
1. Introduction

A currently much debated question in educational research is how actual classroom processes are associated with more general outcomes, such as teacher emotions and burnout (e.g., Frenzel, 2014). Although most theories describe emotions as dynamic intra-individual processes (Frijda, 1986; Scherer, 2009), many empirical studies only used inter-individual (i.e., between-person) comparisons at one moment in time to investigate them. On a general inter-individual level, the interpersonal relationship between teacher and students has been found to be an important factor in the emergence of negative emotions and teacher burnout (Chang, 2009; Hoglund, Klingle, & Hosan, 2015; Roorda, Koomen, Spilt, & Oort, 2011; Spilt, Koomen, & Thijs, 2011; Van Droogenbroeck, Spruyt, & Vanroelen, 2014). However, these findings at an inter-individual level might not generalize to the intra-individual level (Fisher, Medaglia, & Jeronimus, 2018; Murayama et al., 2017), while interventions with teachers most likely need to focus on what happens from moment-to-moment in class (Pennings & Mainhard, 2016; Van Vondel, Steenbeek, Van Dijk, & Van Geert, 2017). The current paper illustrates the use of two process measures capturing teaching as it occurs in terms of a teachers’ physiology and interpersonal behavior.

Both physiology and interpersonal behavior have been described as theoretical components of emotional processes (Scherer, 2009). New methodological developments now enable us to study interpersonal and affective variables as processes at the intra-individual level during real-class teaching. More specifically, the current study explores the value of observational, quantitative time-series data of interpersonal teacher behavior coded from moment-to-moment (Lizdek, Sadler, Woody, Ethier, & Malet, 2012; Sadler, Ethier, Gunn, Duong, & Woody, 2009), combined with continuous heart rate data as a physiological proxy for teachers’ affective processes during teaching (cf. De Geus, Willemsen, Klaver, & Van Doornen, 1995). The aim of the current paper was to illustrate the promises and challenges of such state-of-the-art methods and their analysis by an in-depth quantitative examination of two contrasting cases: Two teachers with pronounced differences in their self-reported habitual discrete emotions and burnout levels.

1.1 Measuring emotions

Emotions are an integral part of the classroom setting. Many studies in academic settings have focused on discrete emotions of students (such as enjoyment, anxiety, and boredom), and their relation with motivation and achievement (e.g., Goetz, Frenzel, Pekrun, Hall, & Lüdtke, 2007; Pekrun, Goetz, Titz, & Perry, 2002). Currently, research on teacher emotions is becoming more and more prominent. Teacher emotions have been associated with teacher wellbeing (Spilt et al., 2011), the quality of teacher-student interaction (Hagenauer, Hascher, & Volet, 2015), as well as with student outcomes in the affective domain (e.g., enjoyment; Frenzel, Goetz, Lüdtke, Pekrun, & Sutton, 2009; Mainhard, Oudman, Hornstra, Bosker, & Goetz, 2018) and the cognitive domain (e.g., report card grades; Reyes, Brackett, Rivers, White, & Salovey, 2012).

Measuring a person’s emotions is challenging (Mauss & Robinson, 2009), and emotion researchers have used a variety of definitions, approaches, and instruments. Kleinginna and Kleinginna (1981) compiled a widely used definition based on 102 different definitions used in research on emotions: “Emotion is a complex set of interactions among subjective and objective factors, mediated by neural/hormonal systems, which can (a) give rise to affective experiences such as feelings of arousal, pleasure/displeasure; (b) generate cognitive processes such as emotionally relevant perceptual effects, appraisals, labeling processes; (c) activate widespread physiological adjustments to the arousing conditions; and (d) lead to behavior that is often, but not always, expressive, goal-directed, and adaptive” (p. 355). Scherer (1999) emphasized that emotions emerge based on an individual’s appraisal of events. However, appraisals are only one component in the multi-componential perspective of Sutton and Wheatley (2003), who define “appraisal, subjective experience, physiological change, emotional expression, and action tendencies” (p. 329) as the components of emotions.

In line with key theorists (e.g. Frijda, 1986; Lazarus, 1991), most studies on emotions in academic settings, and on teacher emotions specifically, have adopted this multi-componential view. These studies often measured the affective, cognitive, motivational, and physiological components of discrete emotions with questionnaires (e.g., Frenzel et al., 2016; Pekrun, Goetz, Frenzel, Barchfeld, & Perry, 2011). Questionnaires
are helpful tools when investigating discrete emotions during salient episodes (Becker, Goetz, Morger, & Ranellucci, 2014; Goetz et al., 2015), but retrospective bias may occur when assessing emotions after the situation of interest has already ended (Wilhelm & Grossman, 2010). In addition, these retrospective questionnaires, that are administered once and typically at the end of a day (or longer period), do not capture fluctuations in teachers’ emotions (Sutton & Wheatley, 2003), whereas current emotion researchers emphasize the value of measuring emotion as a constantly changing process within individuals (Goetz et al., 2015; Hollenstein, 2015; Kuppens, 2015; Scherer, 2009). Indeed, recent findings indicate that important affective differences between people may not lie in mean levels of emotions, but in how people’s emotions are changing from moment-to-moment (Becker, Keller, Goetz, Frenzel, & Taxer, 2015; Hamaker, 2012). Several indicators can be used to assess these affective dynamics, such as variability (i.e., deviation from an individual’s average; Wichers, Wigman, & Myin-Germeyns, 2015), and inflexibility or inertia (i.e., resistance to change/temporal dependency; Hollenstein, 2015; Trull, Lane, Koval, & Ebner-Priemer, 2015). For example, and emphasizing the need to study emotions as processes, increased variability and higher levels of inertia have been related to less successful coping and emotion regulation (Bonanno & Burton, 2013), more depressive symptoms (Koval, Pe, Meers, & Kuppens, 2013), and low psychological well-being (Houben, Van den Noortgate, & Kuppens, 2015).

A recent development in the field of emotion research that taps into these continuous, ongoing aspects of emotions, is the experience sampling methodology (ESM; Csikszentmihalyi & Larson, 1987; Scollon, Kim-Prieto, & Diener, 2009). This method enables recording of (discrete) feelings and activities, by repeatedly asking people to answer a small set of questions in situ. However, repeatedly reporting on emotions might influence subsequent emotions and patterns of emotions over time (Kassam & Mendes, 2013; Scollon et al., 2009; Wilhelm & Grossman, 2010), and places a heavy demand on participants (Scollon et al., 2009). Although ESM has successfully been applied in studies on teacher emotions, it also interrupts the ongoing lesson and therefore endangers the authenticity of the measurement (Becker et al., 2015). In addition, although ESM tackles the issue of retrospective biases, self-report data is sensitive to other biases, such as social desirability (Hagenaer et al., 2015). To overcome these biases, it has been argued that research on teacher emotions may benefit from using physiological measures (Frenzel, Becker-Kurz, Pekrun, & Goetz, 2015).

1.2 Physiological measures to assess emotion

Physiology refers to the study of the structure and function of bodily systems. Psychophysiology, more specifically, is concerned with linking the bodily system to the experience, feelings, and behavior of people (Cacioppo, Tassinary, & Berntson, 2017). While it is evident that many physiological responses are associated with emotions, there is no consensus in the field of psychophysiology on the exact linkage between physiological signals and discrete emotions; physiological measures are typically indirect (Cacioppo & Tassinary, 1990; Gratton & Fabiani, 2016). There have been efforts to tap into discrete emotions by combining several physiological signals (Wilhelm, Pfaltz, & Grossman, 2006; for a review, see Kreibig, 2010), and although some patterns emerged, overall, clear physiological distinctions between emotions are hard to establish. Because of this current state of affairs, in the present paper, we conceptualized a physiological reaction (after correction for physical activity) as an affective response.

During the last decades, technical innovations have produced several devices that are suitable for ambulatory assessment of physiological data (Wilhelm & Grossman, 2010). This enables researchers to measure physiological activation as a proxy for affective responses within authentic contexts like classrooms, thereby increasing the ecological validity of the measurements (Trull & Ebner-Priemer, 2013; Wilson, 1992). Until now, ambulatory devices have mostly been used to compare physiological signals in an aggregated way across two or more conditions (e.g., Kreibig, Gendolla, & Scherer, 2012) and to compare individual mean levels (e.g., Pieper, Brosschot, Van der Leeden, & Thayer, 2010). In the present study however, our interest lies in moment-to-moment changes in physiological activation.

Two bodily systems that are regularly examined in psychophysiological research are the autonomic nervous system (ANS) and the hypothalamic-pituitary-adrenal (HPA) axis. The hypothalamic-pituitary-adrenal axis is a slowly reacting system, that comes into action when the stressor is large or endures for a longer time (Sapolsky, Romero, & Munck, 2000). The autonomic nervous system is the system that reacts
very fast to changes and stressors in the environment (Nederhof, Marceau, Shirkiff, Hastings, & Oldehinkel, 2014), and is therefore more suited for research on emotion dynamics at smaller timescales (i.e., minutes or seconds). The autonomic nervous system consists of the sympathetic nervous system (SNS; ‘fight-or-flight’ system) and the parasympathetic nervous system (PNS; ‘rest-and-digest’ system). In reaction to a stressor, the parasympathetic nervous system puts its actions on hold to support the sympathetic nervous system in the stress response. The sympathetic and parasympathetic nervous system work together to reach an equilibrium during and after a stressful situation (Nederhof et al., 2014).

There are several indices of autonomic nervous system activation, such as cardiovascular activity (blood circulatory system) and electrodermal activity (sweat glands; Mauss & Robinson, 2009; Wilhelm, Grossman, & Müller, 2012). Electrodermal activity (EDA) captures the flow of electricity through the skin of a person. Sweat production increases under stress, resulting in increased conductance of the skin (Sharma & Gedeon, 2012). However, measures of electrodermal activity become less precise when hand palms are too wet and when there are environmental temperature changes (Boucein et al., 2012), or when an emotional trigger is relatively short (Oliveira-Silva & Gonçalves, 2011). The cardiovascular system has been featured in many studies, because it is one of the strongest human biosignals and sensitive to almost all emotions (Myrtek, 2004) besides various other physical and psychological processes (Berntson, Quigley, Norman, & Lozano, 2016). Several cardiovascular measures can be used, such as heart rate (HR), heart rate variability (HRV), blood pressure (BP), and pre-ejection period (PEP; Mauss & Robinson, 2009). Heart rate is one of the most widely used biosignals (Kreibig, 2010), encompasses both sympathetic and parasympathetic activation (Nederhof et al., 2014), and can be measured from moment-to-moment non-invasively (Van Dijk et al., 2013). Therefore, the focus of the present study was on continuous measurement of heart rate as a proxy for an affective response.

Heart rate is usually converted to beats per minute (bpm), but can also be expressed as heart period, representing the time in milliseconds between two heartbeats (Berntson et al., 2016). When applying heart rate measures in a larger sample, it is advised to use heart period, because this measure is less affected by inter-individual variability in baseline heart rate (see Berntson, Cacioppo, & Quigley, 1995; Porges & Byrne, 1992). Heart rate variability (HRV) is another often used measure (e.g., Dishman et al., 2000; Koval, Ogrinz, et al., 2013), and represents the variability of heart period over time. Some drawbacks that limit the potential of heart rate variability for between- and intra-individual comparisons are that it is a purely parasympathetic measure (Mauss & Robinson, 2009), that it needs to be summarized over longer periods of time, and that it is influenced to a relatively large extent by heart rate, age, and respiration (Berntson et al., 2016; Myrtek, 2004). Further, blood pressure has often been used in ambulatory monitoring studies, but also not on a moment-to-moment basis (Holt-Lunstad, Uchino, Smith, Olson-Cerny, & Nealey-Moore, 2003; Vrijkotte, Van Doornen, & De Geus, 2000). There have been some efforts to measure blood pressure noninvasively (Friedman, Christie, Sargent, & Weaver, 2004), but to the best of our knowledge, these methods are only applicable in lab settings due to the large size of the devices. Finally, pre-ejection period is an index of sympathetic nervous system activity only, and is very sensitive to changes in posture or physical activity (Vrijkotte, Van Doornen, & De Geus, 2004). This limits its use for moment-to-moment measures, because movement cannot be restricted in most real-life activities like teaching.

Heart rate is of course always related to physical activity and postural changes (e.g., in order to provide muscles with oxygen), which needs to be taken into account in real-life situations (Houtveen & De Geus, 2009; Myrtek, 2004). According to Rohrbaugh (2016), physiological changes cannot be interpreted without information on the demands placed on the bodily system by physical activity. Therefore, some researchers restricted their analyses to periods with comparable posture and levels of physical activity, which is of course not preferable when studying dynamic processes in real-life settings with many potential changes in posture and movement. Myrtek (2004) describes the Additional Heart Rate (AHR) method, which mathematically controls for movement; an increase in heart rate is continuously compared to the heart rate in a pre-defined preceding period, while controlling for physical activity. In the present study, an affective response is assumed, when the heart rate rises to a larger extent than can be accounted for by physical activity (Carroll, Phillips, & Balanos, 2009; Myrtek, 2004). We thereby view an increase in heart rate (controlled for physical activity) as indicating a change in the environment that is emotionally meaningful to the teacher (cf. Mauss & Robinson, 2009).
1.3 Interpersonal behavior as a context for affective responses

Emotions, including physiological responses, are intertwined with the social context in which they are experienced (Fischer & Van Kleef, 2010). People are continually adapting their emotions based on the interactions and emotional displays of people around them (Butler, 2015; Mainhard et al., 2018; Van Kleef, 2009). Because of this dynamic, in which emotions are affected by and simultaneously affect social interaction and behavior (Haines et al., 2016; Keltner & Haidt, 1999), we investigated teachers’ affective responses in combination with their interpersonal behavior in the classroom context.

Interpersonal theory is one of the most basic approaches to describing the interpersonal meaning of behavior exhibited in the vicinity of others (Horowitz & Strack, 2010). Interpersonal theory posits that two general dimensions underlie all behavior of people in interaction with others: Agency (i.e., dominance, power or social influence) and Communion (i.e., friendliness, affection or warmth). These dimensions are at the same time necessary and sufficient to exhaustively capture interpersonal processes, that is, all behavior displayed in the presence of others can be described as a specific combination of these two dimensions (for empirical evidence, see (Sadler et al., 2009; Wubbels, Brekelmans, Den Brok, & Van Tartwijk, 2006). The Interpersonal Circle for Teachers (IPC-T; see Figure 1; Wubbels et al., 2012) is an adaptation of interpersonal theory to the classroom setting.

![Interpersonal Circle for Teachers](image)

*Figure 1. The Interpersonal Circle for Teachers (IPC-T; Wubbels et al., 2012).*

The IPC-T can be used to assess both self- and student-perceptions of teachers’ general Agency and Communion in class via questionnaires (i.e., Questionnaire on Teacher Interaction, QTI; Wubbels, Créton, & Hooymayers, 1985). Research has shown that teacher behavior characterized by relatively high levels of Agency and Communion, that is, directing and helpful behavior, is preferred by both teachers and students, and is associated with more pleasurable student emotions (Mainhard et al., 2018) and better cognitive outcomes (Roorda, Jak, Zee, Oort, & Koomen, 2017; Wubbels, Brekelmans, Mainhard, Den Brok, & Van Tartwijk, 2016). Sadler et al. (2009) developed a computer-joystick method to observe interpersonal behavior continuously over time: Continuous Assessment of Interpersonal Dynamics (CAID). With this method, teachers’ Agency and Communion can be coded by external observers as it unfolds over time (Pennings, Brekelmans, et al., 2014). Research on these moment-to-moment processes has shown that teachers’ interpersonal behavior is not constant but dynamic, and more variability in teacher interpersonal behavior tends to occur in classrooms with poorer social climates (Mainhard, Pennings, Wubbels, & Brekelmans, 2012; Pennings et al., 2017). The CAID method is described in more detail in section 2.3.2.
1.4 The present study

The aim of the present study was to showcase different approaches to the analysis of physiological and behavioral time-series data in the context of teacher-emotion research. We illustrate both univariate (i.e., separate for each time series) and multivariate (i.e., combining time series) process analyses to investigate the dynamics of teachers’ physiological responses and interpersonal behavior during teaching. Combining continuous physiological measures with continuous coding of interpersonal behavior has the potential to advance our understanding of why teachers vary in their teaching-related emotions. To get an idea of the potential value of process measures for the study of teacher emotion in connection to work-related stress and well-being, we selected two teachers representing contrasting cases (i.e., similar interpersonal student perceptions, but diverging emotion and burnout scores). The teachers’ physiological activation was measured as a proxy for their affective response during a regular classroom lesson, which was also filmed to code teachers’ interpersonal behavior.

2. Method

2.1 Participants

We present an in-depth investigation of process measures of physiology (i.e., heart rate) and interpersonal behavior of two Dutch secondary school teachers from a pilot study ($N = 10$) that was part of the Dynamics of Emotional Processes in Teachers (DEPTh) project. Teacher A was a younger female teacher of social studies with 5 years of teaching experience and teacher B was a somewhat older male science teacher with 20 years of experience. Both classes were of approximately equal size (about 25 students) and consisted of an approximately equal number of boys and girls with a mean age of about 14 years. The selected teachers differed in terms of reported levels of teaching-related discrete emotions (i.e., anxiety and relaxation) and work-related burnout, whereas their general interpersonal teaching style exhibited in class was rather similar according to student questionnaires. An overview of the measures used to characterize the two teachers is provided in Table 1. The dissimilar combination of interpersonal and emotional variables was chosen to be able to explore the predictive potential of intra-individual processes and the potential link with teachers’ emotions reported after the lesson.
Table 1

Description of the Selected Teachers in Terms of Emotion, Burnout, and Interpersonal Style

<table>
<thead>
<tr>
<th>Construct</th>
<th>Teacher A</th>
<th>Teacher B</th>
<th>Population mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Emotion</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relaxation e.g., During this lesson, I was relaxed</td>
<td>1.80</td>
<td>3.60</td>
<td>3.87 (0.97)</td>
</tr>
<tr>
<td>Anxiety e.g., During this lesson, I felt tense</td>
<td>4.00</td>
<td>1.75</td>
<td>1.64 (0.94)</td>
</tr>
<tr>
<td><strong>Burnout</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotional Exhaustion e.g., I feel burned out from my work</td>
<td>1.88</td>
<td>1.25</td>
<td>2.04 (1.20)</td>
</tr>
<tr>
<td>Depersonalization e.g., I do not really care what happens to my students</td>
<td>1.83</td>
<td>0.83</td>
<td>1.63 (0.87)/1.16 (0.67)</td>
</tr>
<tr>
<td>Personal Accomplishment e.g., I deal very effectively with the problems of my students</td>
<td>4.29</td>
<td>5.00</td>
<td>3.83 (0.83)</td>
</tr>
<tr>
<td><strong>General interpersonal style</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agency i.e., teacher dominance</td>
<td>.21</td>
<td>.27</td>
<td>.19 (.24)</td>
</tr>
<tr>
<td>Communion i.e., teacher friendliness</td>
<td>.33</td>
<td>.26</td>
<td>.29 (.34)</td>
</tr>
</tbody>
</table>

* Dutch translation and adaptation of the Achievement Emotions Questionnaire (AEQ: Pekrun et al., 2011) and Teacher Emotions Scales (TES; Frenzel et al., 2016). Means are based on the current sample (N=10).

* Utrecht Burnout Scale (Schaufeli & Van Dierendonck, 2000). Population mean scores are based on a sample of teachers in secondary education (N=603). For Depersonalization, the norm scores are different for males/females.

* Questionnaire on Teacher Interaction (Wubbels et al., 2006, 1985). Population mean scores are based on a sample of 1668 students (Mainhard et al., 2018).

2.2 Procedure

Teachers selected the, according to them, most challenging group of students they currently taught and a lesson with a considerable amount of teacher-student interaction (i.e., no extended seat work). An online questionnaire was administered before the selected lesson, assessing teachers’ personal ideal levels of Agency and Communion in class, self-efficacy, general teaching-related emotions, burnout, and work engagement. Just before the lesson, a heart rate measurement device was attached to the teacher’s chest. Lessons were filmed from the back of the classroom, focused on the teacher. During the last ten minutes of the lesson, both the teacher and the students filled out a paper-and-pencil questionnaire focused on their emotions during the lesson.

The general set-up of this study was approved by the Ethics Committee of the Faculty of Social and Behavioral Sciences of Utrecht University (FETC16-074). All teachers and students included in this study provided their active written informed consent. Students that did not agree to participate were either not videotaped or their faces were made indistinct, and their questionnaire data was not used in the analyses.

2.3 Measures

2.3.1 Heart rate

Teacher’s cardiovascular and physical activity was measured continuously with the VU University – Ambulatory Monitoring System (VU-AMS; www.vu-ams.nl). Seven electrodes were placed on the teacher’s chest to minimize the influence of movement (Porges & Byrne, 1992) and to get a more reliable estimate of heart rate than with wristbands based on pulse plethysmography (PPG; Rohrbaugh, 2016; Schäfer & Vagedes, 2013). The electrocardiogram (ECG) data was processed in three stages: signal enhancement, data reduction, and statistical analysis (following the guidelines of Gratton & Fabiani, 2016). To enhance the signal, we checked the raw data for artefacts (i.e., bad ECG quality) and outliers (i.e., anomalous interbeat intervals, wrongly placed R-peaks) with the help of the automated detection of the VU-AMS software. Data correction or deletion was not necessary for the fragments in the current study. We controlled for the influence of physical...
activity in line with Myrtek’s (2004) Additional Heart Rate approach. A challenge in our case was that physical activity was related not only to heart rate, but also to teachers’ interpersonal behavior (e.g., interpersonal Agency or dominant teacher behavior was associated with relatively more movement). Thus, the part of the physical activity that showed unique overlap with the heart rate signal was filtered out by 1) regressing physical activity on interpersonal behavior, and 2) regressing heart rate on the residual physical activity. The resulting residual heart rate values were used in the process analyses.

In the data reduction stage, one of the main challenges of dynamic process analyses (as compared to averaging over situations or experimental conditions) is choosing a time interval to represent the data (Hamaker & Wichers, 2017). We followed Gratton and Fabiani (2016) in selecting a time period that balanced the richness of the data, which we wanted to maximize, and the amount of noise or measurement error, which we wanted to minimize (Porges & Byrne, 1992). In our case, we used 5 second-averages, as this time interval has at least 4-5 heart beats for each person, but still reflected the small-scale changes we were interested in and which would be lost when applying larger time intervals (e.g., there is no change if data is aggregated into a one-minute aggregate; see Figure 2).

![Figure 2. Heart rate at different levels of aggregation.](image)

2.3.2 Interpersonal teacher behavior

Two trained coders independently rated teacher’s interpersonal behavior continuously according to the CAID approach (Lizdek et al., 2012; Pennings, Brekelmans, et al., 2014; Sadler et al., 2009). This approach enables observers to take into account both verbal and non-verbal behavior with a clear interpersonal meaning (Markey, Lowmaster, & Eichler, 2010). While watching the classroom video on one side of a computer screen, teacher’s interpersonal behavior was assessed from moment-to-moment by moving with a joystick apparatus over the interpersonal circle (see Figure 1) on the other side of the screen. That is, the observers followed the teacher’s behavior and positioned it on the orthogonal Agency and Communion axes simultaneously. The position of the joystick on the axes represents the type of behavior, whereas the distance from the origin embodies the intensity. An increase in Agency was registered, for example, when teachers were directing the conversation vs. following what students demanded. Examples of behaviors that were coded as an increase in Communion were smiling or supporting students, and a decrease would have been recorded when teachers were not responding to students or made ‘cold’ comments (Ross et al., 2017). The software saves separate values for teacher Agency and Communion every half-second, on a scale from minus 1000 to 1000. Intra-class correlations (consistency) between both coders’ ratings were ICC(2,2) = .95 for Agency and ICC(2,2) = .55 for Communion, which can be interpreted as respectively very strong and moderate agreement (Cicchetti, 1994; Koo & Li, 2016; Lebreton & Senter, 2008). For the analyses, the ratings of both coders were averaged to
account for idiosyncratic observations. Following Thijs, Koomen, Roorda, and Ten Hagen (2011), and to match the heart rate data, 5-second averages of Agency and Communion were used in the analyses. We calculated the 5-second average as an aggregate of the ten underlying half-second codes instead of using only one code per five seconds. With this approach, the averages were controlled for possible outliers that could otherwise have influenced the results. The correlation between the resulting values in both approaches were all above .97 (i.e., for both teachers, and for both Agency and Communion).1

2.4 Data-analysis

We focused our analysis on the first 12 minutes of the lesson, because the lesson start is at the same time very important and very demanding regarding teacher-student interaction (Pennings et al., 2017; Van Tartwijk, Brekelmans, Wubbels, Fisher, & Fraser, 1998). First, we examined univariate characteristics of the heart rate, Agency, and Communion time series: a) linear, quadratic, and cubic trend models, which indicate whether teacher’s heart rate (or Agency, or Communion) in- or decreased continuously during the lesson and whether it de- or accelerated at some points, b) autocorrelation, which indicates to what extent heart rate values carry over to the next moment in time (i.e., temporal dependency), and c) stability, which combines the temporal dependency with the amplitude of heart rate changes. These measures provided not only a description of the time series but are also potential predictors of macro-level outcomes. For example, Houben et al. (2015) found in their meta-analysis that higher autocorrelation and lower stability in emotions were associated with lower psychological wellbeing.

Second, the time series of each teacher were analyzed simultaneously in a multivariate fashion to investigate how they were related to each other over time with a cross-correlation analysis (e.g., the extent to which teacher Agency was on average related to heart rate), and by using Dynamic Structural Equation Modeling (DSEM; Asparouhov, Hamaker, & Muthén, 2017) to get insight in sequential or quasi-causal relations between the three time series within a teacher (e.g., whether a decrease in teacher Communion was followed by an increase in Agency, heart rate, or both). These multivariate process analyses of intra-individual dynamics provide information about the coupling of interpersonal and affective processes as they unfold over time. The nature of this coupling, again, represents a potential predictor of macro-level outcomes such as general feelings of burnout (cf. Hamaker, Asparouhov, Brose, Schmiedek, & Muthén, 2018).

One issue in time-series analyses is the handling of trends. In our case we chose not to remove trends, but rather analyze them, as they may capture valuable and essential information, and removing them might blur the interpretation of the results (Boker, Rotondo, Xu, & King, 2002; but also see Warner, 1998; Wu, Huang, Long, & Peng, 2007 for a discussion of detrending time-series data). Data and syntax for this article can be found at https://osf.io/p9ak2/.

3. Results

3.1 Case description

Figure 3 presents the time series of heart rate, Agency, and Communion for Teacher A and B. Each time series consisted of 144 data points (12 minutes x 5-second intervals). In line with the general student perceptions based on the QTI, and as could be expected based on research relating general interpersonal student perceptions with coded interpersonal behavior (Pennings, Van Tartwijk, et al., 2014), the mean level of observed interpersonal behavior, for teacher A and B respectively, was similar for both the Agency ($M = 422$, $SD = 328$; $M = 569$, $SD = 276$) and the Communion dimension ($M = 380$, $SD = 166$; $M = 336$, $SD = 185$). For both teachers, the lesson started with approximately five minutes (300s) where students entered the classroom, and the teacher was talking to a smaller group of students, while other students were talking to each other. After these five minutes, teacher A started the lesson with an overview of the lesson, and discussed a previous

---

1 Results available from the first author upon request
written assessment including a whole-class review of the test. Also, teacher B started with a general review of a test, instructing students with lower scores, followed by an introduction to a group assignment (i.e., to set up a small physics experiment) that students worked on during the remainder of the lesson. The time-series data (see Figure 3) and large SD values showed that the two teachers were variable in their heart rate as well as their interpersonal behavior over time during the lesson start, which indicated that a detailed process analyses could potentially add information beyond examining mean-level associations.

![Figure 3. Time series for heart rate, Agency, and Communion.](image)

### 3.2 Univariate process analyses

#### 3.2.1 Trend

The trend gives an indication of the overall directionality of variables. We explored the existence of linear, quadratic, and cubic trends in the three time series. Table 2 shows that the time series for Agency and Communion for both teachers were best described by the cubic trend model (i.e., linear + quadratic + cubic term), which corresponds to previous findings (Pennings et al., 2017). Halfway the 12 minutes (approximately 300s), there was an increase in Agency for both teachers, reflecting the moment where they took the lead in opening the lesson, and their Agency-level stabilized at about 7.5 minutes (450s). Regarding Communion, teacher A first became friendlier, but then her friendliness decreased while her Agency increased (i.e., she became more imposing and strict). For teacher B the general trend was less pronounced, as the lower amount of variance explained by the trends indicated.
Heart rate showed a significant linear + quadratic trend for Teacher A. Her heart rate first increased for about eight minutes (480s), and then leveled off slightly. This may indicate an affective response in particular to opening the lesson. The heart rate of teacher B showed a somewhat reversed pattern. After a relatively higher starting level, there was a linear decreasing trend for Teacher B, which may be a sign of more effective emotion regulation, that is, the parasympathetic part of the autonomic nervous system might have become more active, resulting in a lower heart rate (Gross, 1998; Sack, Hopper, & Lamprecht, 2004).

Table 2

The Amount of Explained Variance ($R^2$) per Time Series by Adding a Linear, Quadratic, and Cubic Trend

<table>
<thead>
<tr>
<th></th>
<th>Heart rate</th>
<th></th>
<th>Agency</th>
<th></th>
<th>Communion</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Teacher A</td>
<td>Teacher B</td>
<td>Teacher A</td>
<td>Teacher B</td>
<td>Teacher A</td>
<td>Teacher B</td>
</tr>
<tr>
<td>Linear</td>
<td>.15**</td>
<td>.16**</td>
<td>.67**</td>
<td>.67**</td>
<td>.29**</td>
<td>.03*</td>
</tr>
<tr>
<td>+Quadratic</td>
<td>.26**</td>
<td>.16**</td>
<td>.73**</td>
<td>.68**</td>
<td>.63**</td>
<td>.30**</td>
</tr>
<tr>
<td>+Cubic</td>
<td>.27**</td>
<td>.17**</td>
<td>.85**</td>
<td>.82**</td>
<td>.66**</td>
<td>.38**</td>
</tr>
</tbody>
</table>

* p < .05
** p < .001

3.2.2 Autocorrelation

Autocorrelations indicate how well one state in a time series can be predicted by its previous state, this is also referred to as carry-over effect, inertia, or temporal dependency (Houben et al., 2015). A high autocorrelation value on affective variables is usually interpreted as low reactivity to changes in the environment and less recovery, and this has been related to lower psychological wellbeing and psychopathology (Houben et al., 2015; Wichers et al., 2015). For interpersonal behavior, a high autocorrelation appears to be associated with more positive teacher-student relations (i.e., predictable behavior and few abrupt changes; Mainhard et al., 2012; Pennings, Brekelmans, et al., 2014).

In the present study, the autocorrelations for Teacher A and B respectively were .47 vs. .27 for heart rate, .99 vs. .98 for Agency, and .92 vs. .94 for Communion. Thus, the autocorrelations indicated that there was a higher carry-over effect of heart rate for Teacher A. Following Hollenstein (2015), this can be seen as Teacher A being less flexible in her affective response and potentially less adaptive to specific classroom situations, or being less able to regulate her affective response (e.g., less cognitive reappraisal; Gross & John, 2003). For both teachers, there was a large carry-over effect for interpersonal behavior, which indicated that both teachers were quite predictable in their behavior (i.e., not many sudden changes from friendly to unfriendly or from dominant to submissive behavior).

We checked whether the high autocorrelations for Agency and Communion could be explained by the small time frame between two data points (i.e., whether interpersonal behavior might take more than five seconds to change substantially) by testing larger time intervals and by using a lag of two time points. Exploratory analyses with larger time intervals (i.e., 10 and 30 seconds) showed that interpersonal behavior still had a relatively large autocorrelation. The autocorrelation became negative when leaving out the first time lag and using a time lag of two, which might indicate a cyclical pattern. However, as research has shown that trends and cyclical patterns are individual teacher characteristics (Pennings et al., 2017), and in the absence of strong methodological guidelines (Hamaker & Wichers, 2017), we did not want to draw conclusions based on the data of only two teachers, but decided to follow previous studies in choosing a first-order time lag and a relatively small time interval (Dormann & Griffin, 2015). Moreover, although both teachers’ level of interpersonal behavior was highly predictable from the value five seconds before, this does not mean that there were no meaningful fluctuations in Agency and Communion during the lesson start (as is also evident from the graphs presented in Figure 3).

3.2.3 Stability

Stability combines the temporal dependency with the variability of time series. This measure illustrates not only whether participants are changing (i.e., auto-correlation), but also takes into account how large these

---

2 Results available from the first author upon request
changes are (e.g., standard deviation). In the teaching context, small changes from moment-to-moment might be seen as adaptive; however, large changes (instability) may make the teachers’ behavior unpredictable for students and may primarily indicate changes towards more negative teacher behavior (Mainhard et al., 2012). Also, affective instability has been related to psychological maladjustment (Koval, Ogrinz, et al., 2013).

The mean square successive difference (MSSD; Jahng, Wood, & Trull, 2008) is a stability measure which captures this variability between measurements and is not influenced by trends in the data (Hamaker, Ceulemans, Grasman, & Tuerlinckx, 2015). As the name suggests, MSSD is the average of the squared difference between moment \( t-1 \) and \( t \). Lower values indicate more stability. The MSSD, for teacher A and B respectively, was 152 vs. 44 for heart rate, 1483 vs. 1454 for Agency, and 3517 vs. 3051 for Communion. These results showed that Teacher A was somewhat less stable in all variables compared to Teacher B, in particular regarding heart rate, which might be related to her self-reported overall more negative feelings about the lesson.

Although teacher A was less stable in her heart rate (i.e., higher MSSD), there was a larger carry-over effect (i.e., high autocorrelation) at the same time. In this case, the MSSD might have been influenced mainly by the higher variability (i.e., standard deviation) of the raw heart rate changes in teacher A (Wang, Hamaker, & Bergeman, 2012). Another interpretation might be that teacher A was more rigid in her affective responses, but that changes in affective response (although less frequent) were relatively large and demanding to regulate, for example, when she had to discipline students. Such a pattern would correspond with teacher A reporting less relaxation and more anxiety after the lesson, and might thus be a risk factor for psychological problems (Hollenstein, 2015; Houben et al., 2015; Koval, Ogrinz, et al., 2013; Wichers et al., 2015).

### 3.3 Multivariate process analyses

By combining the time series, we explored teachers’ heart rate in the context of specific interpersonal behavior, that is, we examined the intertwinedness of a teacher’s interpersonal behavior and affective response. Figure 4 visually combines teacher heart rate with a teacher’s behavioral trajectory projected on the IPC-T. The color of the dots represents the teacher’s heart rate at the moment at which this specific combination of Agency and Communion was observed. Overall, both teachers showed comparable behavioral trajectories (i.e., positive levels of Agency and Communion), which corresponded to their similar general interpersonal style as perceived by their students (cf. Pennings, Van Tartwijk, et al., 2014). Figure 4 shows that teacher A had, based on her individual range, a relatively high heart rate during the entire lesson start (i.e., more dark/red dots overall), and more specifically when showing directing behavior (i.e., combining high Agency and moderately high Communion). In contrast, teacher B had a relatively low heart rate while displaying directing behavior. The trajectory of teacher B indicated one specific situation, where this teacher had to correct a student, became less friendly (negative Communion) and had a relatively higher heart rate.
3.3.1 Cross-correlation

When combining time series in analyses, the cross-correlation represents the overall coordination between time series over the entire time frame. As with regular correlation coefficients, this value can be interpreted as a combination of the strength and direction of the concurrent relationship between two variables. For both teachers, the cross-correlations indicated a negative association between Agency and Communion, but more strongly so for Teacher A (see Table 3). Thus, relatively higher levels in Agency tended to go together with relatively lower levels in Communion (and vice versa). This can be interpreted as a potential interpersonal pitfall, as combining high Agency with relatively high Communion can be understood as more productive in class (i.e., directing instead of imposing teacher behavior; Aelterman, Vansteenkiste, Van den Berghe, De Meyer, & Haerens, 2014; Mainhard et al., 2018). Interestingly, for both teachers, changes in affective responses were more strongly related to changes in Agency than Communion. However, this association was different in nature for the two teachers. For teacher A, the association between Agency and affective response was positive (but statistically non-significant), whereas this association was negative for teacher B. Thus, teacher B tended to have an affective response (i.e., relatively higher heart rate) in situations characterized by relatively lower teacher Agency. Again, these patterns may represent potential predictors that help to understand different emotional outcomes for teacher A and B.

Table 3

Cross-Correlations between Teachers’ Affective Response (i.e., heart rate), Agency, and Communion (N=144 time points)

<table>
<thead>
<tr>
<th></th>
<th>Heart rate</th>
<th>Agency</th>
<th>Communion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Teacher A</td>
<td>Teacher B</td>
<td>Teacher A</td>
</tr>
<tr>
<td>Heart rate</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Agency</td>
<td>.11</td>
<td>-.23**</td>
<td>-</td>
</tr>
<tr>
<td>Communion</td>
<td>.01</td>
<td>-.07</td>
<td>-.63**</td>
</tr>
</tbody>
</table>

**p < .001
3.3.2 Dynamic Structural Equation Modeling

Cross-correlations describe the nature of the concurrent association between two time series (i.e., at the same point in time). However, cross-correlation analysis does not take into account relationships over time and (the in our case large) autocorrelations, which may lead to overestimation of the cross-correlations (see Hamaker et al., 2018; Sadler et al., 2009). Dynamic Structural Equation Modeling (DSEM; Asparouhov et al., 2017) is a new technique that gives insight in cross-lagged relations between variables, that is, how one variable is associated with other variables at the previous time point, while controlling for their own value at the previous time point. The DSEM results for Teacher A and B are presented in Figure 5. These indicated a positive sequential relation between heart rate at \( t-1 \) and Agency at \( t \) for teacher A. Thus, her affective response may have resulted from anticipating situations where she had to become more agentic in class. Also, Agency at \( t-1 \) had a negative association with Communion at \( t \) for this teacher. This corresponds to the negative association between Agency and Communion we discussed above (i.e., moving towards imposing or strict behavior); but in addition the cross-lagged effect suggested that less friendly behavior of Teacher A (lower Communion) tended to be preceded by higher Agency and not the other way around. In contrast, for Teacher B, Agency was negatively associated with heart rate at the next time point. This teacher seems to be less aroused after taking a position with relatively more Agency (e.g., he may feel more at ease in situations where he has the lead in class). This might be related to his level of experience in teaching, as beginning teachers often have more difficulty with this type of classroom management (Aldrup, Klusmann, & Lüdtke, 2017; Brekelmans, Wubbels, & Van Tartwijk, 2005).

Figure 5. Dynamic Structural Equation Model (DSEM; Asparouhov et al., 2017). The subscript \( t \) refers to a time point and the subscript \( t-1 \) refers to the previous time point of the time series. The reported autocorrelations are slightly different from section 3.2.2 due to the addition of the cross-lagged associations. Only statically significant cross-lagged effects are presented (\( N=144 \) time points per teacher). Hr = heart rate; Ag = Agency; Com = Communion.

4. Discussion

Teacher-student relations have been associated with teacher outcomes, such as emotions and burnout, on a general level (Chang, 2009; Roorda et al., 2011; Spilt et al., 2011). To better understand how these associations emerge from the classroom situation, and in order to design individualized interventions, there has been a recent call for intra-individual process analysis in educational research (Murayama et al., 2017; Pennings & Mainhard, 2016; Van Vondel et al., 2017). We therefore explored the value of measuring and analyzing teachers’ interpersonal behavior and physiological processes as two important mechanisms in the emergence of more habitual emotions.

The present study showed that in-depth intra-individual analyses of physiology and interpersonal behavior uncovered potentially meaningful teacher-specific characteristics that may go unnoticed when using, for example, more general mean-level indicators. The two teachers examined here showed marked differences
in intra-individual tendencies, which may be related to divergent self-reported emotions and burnout, despite their similar habitual interpersonal styles in class as perceived by their students and comparable mean levels of observed interpersonal behavior.

The use of physiological measures enabled us to get some insight into teachers’ affective responses without disrupting the teaching process (Mauss & Robinson, 2009) and to reduce issues with social desirability, retrospective bias, and high cognitive load (Becker et al., 2015; Goetz et al., 2015; Scollon et al., 2009; Wilhelm & Grossman, 2010). Moreover, we found that heart rate measures discriminated between both teachers, even when their interpersonal behavior during the lesson start was relatively similar.

Especially combining physiological data with teachers’ interpersonal behavior on a moment-to-moment level, which is directly in line with prominent theories on human emotional functioning (Frijda, Kuipers, & Ter Schure, 1989; Scherer, 2009), enabled us to estimate teacher-specific links between physiology and interpersonal behavior (i.e., cross-correlations or cross-lagged effects). This link may represent teachers’ individual action tendencies (i.e., their habitual reaction pattern to stimuli in the classroom environment), which is one important component in the process leading to emotional outcomes (Frijda et al., 1989; Scherer, 2009). Moreover, it might not be the situation itself that leads to emotional outcomes, but rather the individual appraisal of the situation (Chang, 2013; Frenzel, 2014; Moors, Ellsworth, Scherer, & Frijda, 2013). This might explain why similar interpersonal behavior could result in different emotions for the teachers in our study. For example, whether the situation was in line with the teachers’ lesson goals, or whether the teacher felt able to cope with the situation, is likely to have influenced their physiological (affective) and behavioral response (Kreibig et al., 2012). It might be beneficial for teachers and teacher educators to become aware of such (often tacit) appraisal processes and action tendencies, for example through video feedback coaching (e.g., Van Vondel et al., 2017).

Beyond investigating more general patterns of within-lesson physiological and interpersonal processes, the richness of the collected data in the present study would also allow pinpointing the specific classroom situations that trigger these particular affective responses. Heart rate could be used to select moments during the lesson (e.g., selecting episodes with a particularly high heart rate) and teachers could be asked to reflect on the causes and consequences of their behavior (e.g., their appraisals and expectations). This could be a first step towards identifying for example less functional appraisals or emotion regulation strategies (Gross & John, 2003).

4.1 Limitations and future directions

We used a quantitative contrasting cases approach with two teachers differing in their general level of emotions and burnout to gauge the value of intra-individual process measures for explaining inter-individual differences between teachers. Case studies are quite frequent in the emerging field of intensive longitudinal data due to the large time investment needed to collect, code, and analyze the data. However, we agree that the preeminent value of physiological and moment-to-moment data lies in their explanatory value of higher-level outcomes in a larger sample (Hamaker & Wichers, 2017; Martin et al., 2015). These multilevel models are ultimately needed to check our current speculation about the relation between moment-to-moment dynamics of physiology and interpersonal behavior and general-level teacher emotional outcomes. In larger samples it would be important to investigate also how other teacher characteristics that were contrasting in our two teachers, such as gender or experience, are related to both intra-individual dynamics as well as between-person outcomes. DSEM provides the opportunity to model these relations in an integrated analytical framework, i.e., between time-series variables (e.g., autocorrelations and cross-lagged associations) and higher-level predictor or outcome variables (e.g., self-reported emotions), by modelling random slopes that allow for individual differences in the time-series analysis (Asparouhov et al., 2017; Hamaker et al., 2018). Although recruiting teachers for this type of intensive longitudinal studies is challenging, a relatively large sample (N>50) might be needed to test these multilevel models (Schuldtzberg & Muthén, 2018). It should also be taken into account that the DSEM outcomes only hold for the specific time interval tested, and that there might be, as in most statistical analyses, omitted variables that may change the model (Hamaker et al., 2018). Future research should therefore also include other time lags or implement continuous time models (cf. Deboeck & Preacher, 2015).
There are several other avenues for future research. In the current study, we chose to use heart rate as an indicator of teachers’ affective response, because of its reactivity and robustness in terms of signal quality in real-life situations. Other physiological measures (such as electrodermal activity or heart rate variability) might be worthwhile to explore as well, as there might be different relations with interpersonal behavior and emotions (Kreibig, 2010). We used the Additional Heart Rate approach to control the heart rate signal for physical activity before the analyses. However, in a multilevel framework, physical activity could be added directly to the model. Also triangulation with other unobtrusive measures such as automatic face detection might be a possible avenue for further research on teachers’ emotional processes (e.g., Bosch et al., 2015). Furthermore, combining several time-series descriptors and cross-lagged patterns might help to understand their interrelatedness (e.g., by using latent class/mixture modeling; Flunger et al., 2015; Van den Bergh & Vermunt, 2017). Finally, teachers’ interpersonal behavior is only one of the two constituent parts that make up classroom interaction. In future research, it would be interesting to combine physiological measures with observation of students’ interpersonal behavior, to see whether, for example, refraining from unfriendly reactions to hostile student behavior is as challenging for teachers as is often claimed (De Jong, Van Tartwijk, Verloop, Veldman, & Wubbels, 2012; Pennings et al., 2017; Sadler et al., 2009).

4.2 Conclusion

To conclude, we want to emphasize that moment-to-moment behavioral dynamics as well as physiological data have a role to play in gaining understanding of the complex interpersonal and affective processes underlying teaching and emotions. Based on our exploration of two contrasting cases, we believe that differences between teachers in their intra-individual physiological and behavioral processes may reflect the working of teachers’ action tendencies and appraisal processes and that these might be related to teacher emotions and burnout symptoms. Future research may need to combine different physiological indicators and both inter- and intra-individual analyses to sketch a more complete picture of emotionally relevant processes in the classroom (cf. Mauss & Robinson, 2009). In the end, this might help to advance teachers’ interpersonal behavior and affective responses during the lesson, and thereby improve both teachers’ job satisfaction (Chang, 2013; Hakanen, Bakker, & Schaufeli, 2006) and student outcomes (Arens & Morin, 2016; Hoglund et al., 2015).

Keypoints

- Teacher-student relations, teacher emotions, and burnout are related on a general level, but information is missing on underlying classroom processes
- Continuous Assessment of Interpersonal Dynamics is a method to code teachers’ interpersonal behavior during teaching from moment-to-moment
- Heart rate is a non-invasive and objective physiological measure, and can be used as a continuous indicator of teachers’ affective responses
- Teachers differed in their intra-individual physiological and behavioral processes, potentially indicating individual action tendencies and appraisals
- Time-series analyses have the potential to strengthen research on the link between classroom social processes, teacher emotions, and burnout

Acknowledgments

This research was funded by the Netherlands Initiative for Education Research (NRO/PROO grant 405-14-300-039), which resides under the Netherlands Organization for Scientific Research (NWO). The authors would like to thank the teachers and their classes that participated in the study, and the students for their assistance with data collection and coding.
References


Chang, M. L. (2013). Toward a theoretical model to understand teacher emotions and teacher burnout in the


https://doi.org/10.1371/journal.pone.0064959


Schäfer, A., & Vagedes, J. (2013). How accurate is pulse rate variability as an estimate of heart rate


183 | F L R


