***Working with this manuscript:::***

The methods/analysis/results section and supplements of the paper will be written in Rmarkdown which allows the publishing of a fully reproducible manuscript (data and analysis code). In order to write the intro and discussion together in google docs in a way that can be quickly merged with the manuscript, here are some things to consider while writing:

* ***General****:*
  + note that formatting like bold or italics will not be preserved in the markdown manuscript unless its marked with the respective syntax:
    - one \* for italics: \*example\* → *example*
    - two times \* for bold print: \*\*example\*\* → **example**
* ***Headlines****:*
  + All headlines are marked with # (# heading)
  + smaller subheadings are created by adding # (## subheading level 2; ### subheading level 3)
* ***Citations***
  + if you are in a hurry, just past the DOI of the respective paper in brackets. I’ll sort it out later.
  + If you want to be a real sweetheart and add a full citation:
    - go to googlescholar and search the paper
    - click on the citation icon and then select bibtex
    - copy the entire bibtex entry into the googledoc named “manuscript\_bibtex”. it's in the same google-folder as this document
    - copy the citation ID from the bibtex-code (that would be the entry after the first curly bracket and before the first comma. Its usually firstauthor+year+firstwordtitle (e.g. tomasello2010origins)
    - add the citation in text by using the following syntax: [@paperID]
      * e.g....as discussed in previous work [@tomasello2010origins]

**Thanks for your patience!**

**## Present investigation**

The aim was to investigate whether HR measures assessed by a Fitbit Charge 4 are a suitable and effective method to \*\*(1)\*\* map teachers course of arousal over the course of a five phase lab study, including a micro teaching unit,  and \*\*(2)\*\*  examine whether HR measures can be predicted by self-reported data on cognitive appraisal.

Within the time frame of approximately two hours, we investigated five intervals with a duration of 10-minutes each: In the (1) pre-teaching phase, the subjects were prepared for the following teaching lesson and familiarized with the setting. During the (2) teaching phase, the participants taught a 15-minute self-prepared lesson to a "class" of three actors that simulated nine classroom disruptions. In the following (3) post-teaching phase, the subjects answered a questionnaire and in the (4) interview phase the subjects watched a pre-recorded video of their 15-minute lesson to assess the self-report data of how disrupted subjects felt and how confident they felt in dealing with disruptions. In the (5) end phase, the subject answered a second questionnaire.

According to previous findings that fitness trackers can be used as a low-cost, non-invasive method of measuring HR [hajj2022wrist; @fuller2020reliability] and that different HRs of teachers can be measured in different teaching phases [@donker2020associations; @junker2021potential], we formulate the hypotheses as follows:

\*\*Hypothesis 1\*\*. In a first step, we wanted to display exploratively the course of HR during the entire study. Additionally, we presumed the highest mean HR in the (2) teaching phase and lower mean values in all other phases (\*\*Hypothesis 1a\*\*). Moreover, we expected an increase in HR in the (1) pre-teaching phase as the first phase and a decrease in the following phases (\*\*Hypothesis 1b\*\*).

\*\*Hypothesis 2\*\*. We statistically predicted the subjects’ standardized mean HR for the (2) teaching, the (3) post-teaching, the (4) interview and the (5) end phase with teaching experience. We expected a lower HR in teachers with more teaching experience for the four phases (\*\*Hypothesis 2a\*\*). According to the relationship between physiological arousal and cognitive appraisal, we controlled for shared variance with the self-reported data. We expected for the four phases that teachers who reported that they felt disrupted by disruptions would have higher standardized mean HR (\*\*Hypotheses 2b\*\*). In contrast, individuals with high confidence in dealing with disruptions would have a lower standardized mean HR in the four phases (\*\*Hypothesis 2c\*\*). When considering the three predictors in concert and controlling for their common variance, we expected teaching experience and self-report data to remain substantial predictors (\*\*Hypothesis 2d\*\*).

**# Method**

**## Participants**

In total, the sample consisted of *N* = 81 pre- and in-service teachers. The subjects were recruited via personal contact, email lists and flyers. From the originally assessed data of 84 subjects, three participants were excluded from the analyses due to insufficient data quality, yielding the analysis sample of 81.

The subjects (*n* = 52 females, *n* = 29 males) had a mean age of 31 years (*SD* = 10.90; range: 19-60) and an average teaching experience of 5.64 years (*SD* = 9.46; range: 0-37).

14.81/70.37/9.88% of the subjects were (studying to become) teachers for primary/secondary/special educational needs schools.

All study procedures were carried out in accordance with the ethical standards of the University’s Institutional Review Board and the authors received a positive vote on the study procedures from the Ethics Committee. All participants were informed in detail about the aim and intention of the study prior to testing. Participation was voluntary and only took place after written consent had been given.

**## Variables**

**### Heart Rate**

We used a Fitbit Charge 4 to measure the teachers’ HR. In line with the manufacturer's instructions, the device was attached a finger’s width above the subject’s wrist bone [@fitbitnd]. To determine the HR, the Fitbit flashes green LEDs many times per second and uses light-sensitive photodiodes to measure the volume changes in the capillaries and then calculates how many times the heart beats per minute (bpm).

**### Cognitive appraisal**

Cognitive appraisal was assessed by two 11-point rating items, ranging from 0 (not at all) to 10 (extremely). The first item referred to the disruptiveness of the event (How disruptive was this event for you?); the second item referred to the subject's confidence in dealing with the event (How confident did you feel in dealing with this event?).

**## Procedure**

Within a timeframe of approximately two hours, teachers taught a 15-minute self-prepared lesson to an audience of three actors who simulated typical classroom disruptions. Subsequently, the subjects filled out a questionnaire and were interviewed about the previously taught lesson. For analyzing the HR data, we selected five 10-minute intervals of theoretical interest from this two-hours study:

The (1) pre-teaching phase was the first 10-minute interval of interest. After the experimenter welcomed the subject and put on the watch, the procedure of the study was briefly explained and written consent to voluntarily participate in the study was requested. Next, the subject was asked to prepare the necessary materials for the lesson (connecting the laptop to the beamer, preparing worksheets, etc.). Once the preparation was completed, a warm-up phase took place to familiarize the subject with the laboratory setting. This warm-up phase consisted of two parts: In the first part, the subject and the three actors playfully learned each other's names. The second part involved getting into conversation with each other by asking authentic questions that were not tailored to the role of the actors. The preparation time before the lesson started lasted 10-15 minutes on average, depending on different issues (preparing the lesson, technical problems, etc.). Most subjects spent this phase standing and slightly moving during the warm-up game.

The second 10-minute interval was the (2) teaching phase in which the subject taught a self-prepared lesson to the three actors. Before the lesson started, the subject was instructed to behave and move as naturally as possible during the lesson as in a real classroom. In advance, all subjects were given information about planning their lesson in a meeting to ensure an appropriately realistic teaching situation, e.g. no long film clips to ensure the teacher-student-interaction. During the lesson, the actors were instructed to simulate nine typical classroom events, three each in the following categories: (a) verbal disruptive behavior (chatting with the neighbor, whispering, heckling), (b) physical disruptions (clicking with a pen, drumming with hands on the table, snipping) and (c) lack of eagerness to learn (drawing on a sheet of paper, putting the head on the table, looking at the phone). The order of the events and the performing actors were counterbalanced for a full Latin square design. The actors were trained before to perform the disruptions identically in each data collection. After a short familiarization phase for the teacher of two and a half minutes, the instructions appeared as intervals (every 90 seconds for 30 seconds) on a screen that was only visible for the class. The actors were trained to stop the disruptive behavior as soon as the teacher intervened. Time management was regulated by the experimenter by showing time cards of the remaining time. The lesson lasted about 20 minutes on average. The subjects’ position in this phase was mostly standing and moving during the lesson.

In the (3) post-teaching phase as the third 10-minute interval the subject as well as the actors were given a short questionnaire, which contained items to collect demographic information as well as items about the previously given lesson on teaching quality. The completion of the questionnaire took approximately 10 minutes. The majority of subjects completed the questionnaire in a seated position.

The fourth 10-minute interval of interest was the (4) interview phase in which the subject watched the pre-corded video of the lesson. The experimenter stopped the video at the nine disruptions when they appeared to ask several questions that were identical for each disruption in a fixed order. First, the subject was asked to describe the disruption, then to evaluate (11-point rating scale) and justify the disruptiveness of each disruption. Next, the subject was asked to describe and justify the reaction. The experimenter then asked the subject to evaluate (11-point rating scale) and justify the confidence the subject had in dealing with the disruption. Statements about the evaluation of the disruptiveness and the confidence were quantified. For this purpose, subjects determined their individual value on a rating scale from 0 (not disturbing / not confident) to 10 (disturbing / confident). The interview lasted on average 45-60 minutes. The subjects’ position in this phase was seated and calm during the interview.

The fifth interval was the (5) end phase in which the subject answered a second questionnaire mainly seated and in a calm position. The completion of the questionnaire lasted approximately 10-15 minutes. After completing the questionnaire, the Fitbit watch was removed and the subject was dismissed.

**## Data analysis**

We conducted all analysis with R [@RStudio2020]. For analyzing the HR data, we focused on five 10-minute intervals of theoretical interest:

The first interval, the (1) pre-teaching phase, was calculated from the moment the subject arrived and the Fitbit watch was put on, which happened immediately after the subject was welcomed. The second interval, the (2) teaching phase, began with the experimenter noting the time and step count of the fitness tracker. To ensure that the analysis interval starts with the teaching activity, another two minutes were added to the noted time. The (3) post-teaching phase began immediately after the end of the teaching lesson. To calculate this phase, the time had to be noted as well. The (4) interview phase was 10 minutes in the middle of the interview where we calculated the difference from the end of the lesson and from the time when the subject took off the watch. This duration was divided in two to get to the middle of the interval. Then, 5 minutes were subtracted to get to the start of the 10-minute interval. The (5) end phase was calculated by subtracting 10min from the time the subject took off the watch, which was also the end of the study, in order to analyze the last 10min of the study.

We selected 10-minute intervals for multiple reasons: First, 10 minutes was the minimum duration of all intervals, so we ensured the comparability of the intervals for all participants. Second, @lu2008can confirmed in their study that 10-minute intervals are a useful duration for analyzing photoplethysmography (PPG) data. Third, other studies revealed that the first minutes of the lesson start are essential regarding teacher-student interaction [@donker2018quantitative; @claessens2017positive].

To account for individual differences in the baseline HR, unstandardized values in bpm were z-standardized. Thus, resulting values can be interpreted as differences from the overall HR mean in standard deviations.

Regarding the course of HR during the study, we first displayed the anticipated trend for the mean HR in beats per minute and the standardized HR over the course of the entire study.

To test \*\*Hypothesis 1a\*\*, we carried out paired *t*-tests and the effect size *d* [@cohen1988new] to test for differences in standardized HR using HR means between the (2) teaching phase with all other phases.

For testing \*\*Hypothesis 1b\*\*, we conducted a linear estimation of the increase and decrease in HR over time. To this end, we used fixed intercept-fixed slope regression [@gelman2006data] to estimate linear slopes and intercepts for all individuals. Although this procedure does not account for nonmonotonic progressions in individual HR, a graphical evaluation reveals that the linear estimates fit well for most of the cases (see XX in the appendix).

Regarding \*\*Hypothesis 2a\*\*, we used multiple linear regression to investigate the influence of teaching experience on the subjects’ standardized mean HR for the (2) teaching phase, the (3) post-teaching phase, the (4) interview phase and the (5) end phase. We controlled for shared variance with the self-reported data referring to the disruptiveness of the event (\*\*Hypothesis 2b\*\*) and to the subject's confidence in dealing with the event (\*\*Hypothesis 2c\*\*). In a following step, we considered the three predictors in concert and controlling for their common variance (\*\*Hypothesis 2d\*\*).

**# Results**

**## Heart rate course**

Concerning our first research question, we displayed the trend of the course of HR during the entire study in a first step. Teachers’ unstandardized mean HR over the entire course of the study was *M* = 90.09 bpm and *SD* = 15.76 bpm (range 51 - 164 bpm). The standardized mean HR over the entire course of the study was *M* = -0.04 and *SD* = 0.99 (range -4.03 - 4.56).

Means, standard deviations, and the range of teachers’ unstandardized and standardized HR for the different phases are shown in Table XX.

Table XX M*,* SD *and range of teachers’ HR for the individual phases (unstandardized in bpm/standardized)*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Phase | *M* | *SD* | Min | Max |
| (1) Pre-teaching phase | 96.28/0.48 | 14.11/0.88 | 56/-3.56 | 139/3.24 |
| (2) Teaching phase | 100.80/0.85 | 16.23/0.77 | 63/-2.18 | 164/4.37 |
| (3) Post-teaching phase | 93.61/0.27 | 14.01/0.76 | 60/-2.17 | 150/3.06 |
| (4) Interview phase | 82.32/-0.72 | 11.85/0.74 | 51/-2.51 | 132/4.39 |
| (5) End phase | 77.95/-1.07 | 11.14/0.57 | 50/-2.68 | 120/2.96 |

Based on the HR values shown in Table XX, the subjects started with a comparatively high mean HR already in the (1) pre-teaching phase (*Mstandardized* = 0.71; *SDstandardized* = 0.75). The highest mean HR measure was in the (2) teaching phase (*Mstandardized* = 0.71; *SDstandardized* = 0.75) and the lowest mean HR measure was in the (5) end phase (*Mstandardized* = -1.10; *SDstandardized* = 0.57). A decrease in HR for the following phases can be observed after the teaching phase.

Fig. XX a. shows the unstandardized mean HR in bpm and Fig. XX b. shows the standardized mean HR over the entire course of the study of approximately two hours.

Figure XX *Overall Course of the HR in Beats per Minute and Standardized*



*Note:* The shadow around the line represents the 99% confidence interval. We used the ggplot2 package (v3.3.3; Wickham, 2016) to calculate the moving average of the course.

Fig. XX a. and Fig. XX b. clearly show that in reference to the course of the HR, the HR increased in the (2) teaching phase and decreased in the following phases. Comparing both courses, it is apparent that the course of the non-standardized mean HR is similar to the course of the standardized mean HR.

Referring to the \*\*Hypothesis 1a\*\*, the standardized mean HR in the (2) teaching phase compared to the (1) pre-teaching phase was significantly higher, *t*(80) = 4.99, *p* < .05, *d* = 0.82 (large effect). When comparing the standardized mean HR between the (2) teaching phase and the (3) post-teaching phase, the results revealed a statistically significant difference between these two phases, *t*(80) = 7.59, *p* < .05, *d* = 1.34 (large effect) as well as the comparison between (2) teaching and (4) interview phase, *t*(80) = 18.29, *p* < .05, *d* = 3.37 (large effect). The standardized mean HR in the (2) teaching phase compared to the (5) end phase was significantly higher as well, t(80) = 27.94, *p* < .05, *d* = 4.68 (large effect).

Figure XX*1% Confidence Interval in each phase*

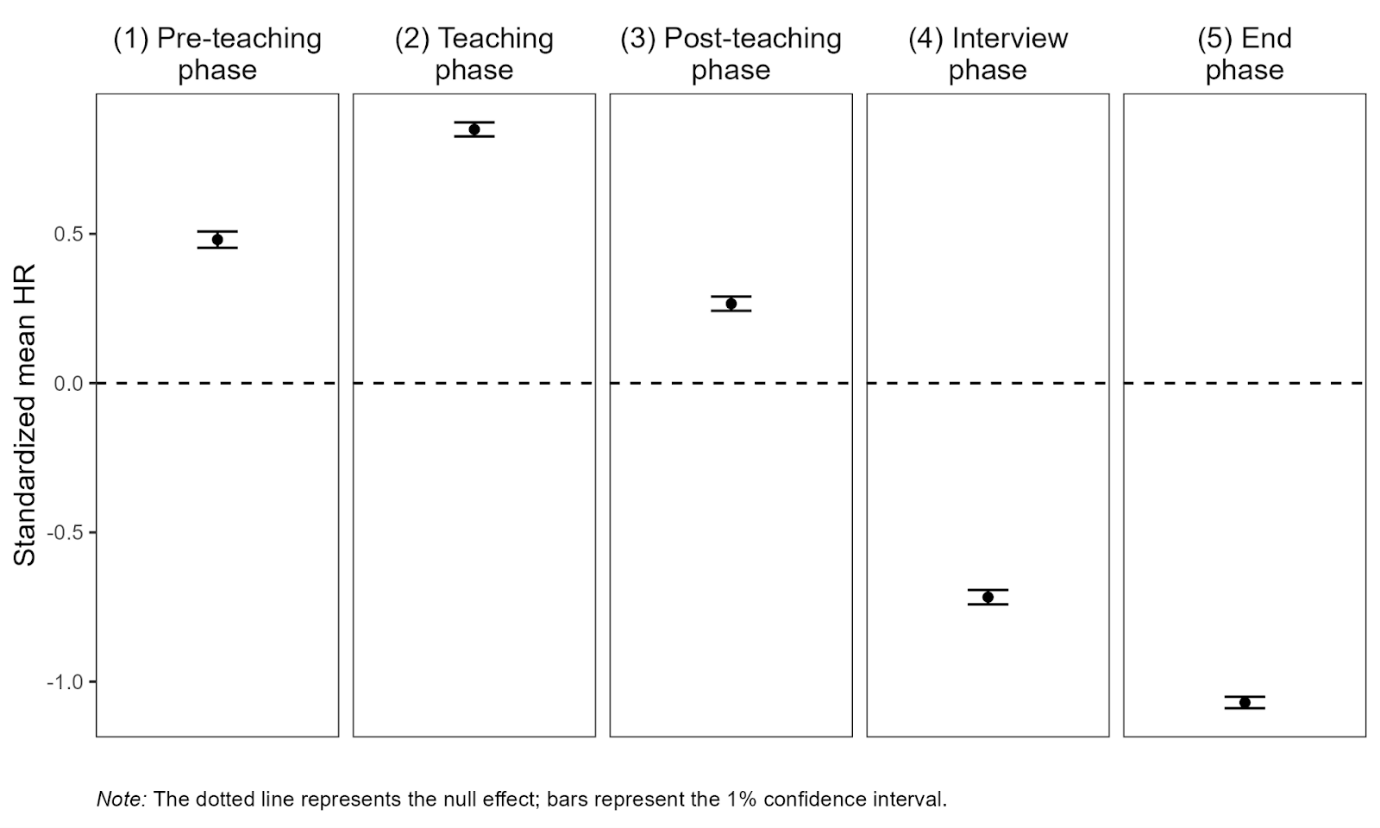
**

Fig. XX shows the mean standardized HR for the five different phases. The trend, which showed up in the entire course of the HR, can be found in the figure as well. We can see the highest mean HR in the (2) teaching phase and lower mean values in all other phases.

Referring to \*\*Hypothesis 1b\*\*, we estimated linear slopes for all individuals to investigate the increase and decrease in HR over time. Although this procedure does not account for nonmonotonic progressions in individual HR, a graphical evaluation reveals that the linear estimates fit well for most of the cases (see … in the appendix). Additionally, we calculated the intercepts.

Table XX shows the descriptive statistics for the mean intercepts and the mean slopes for the different phases.

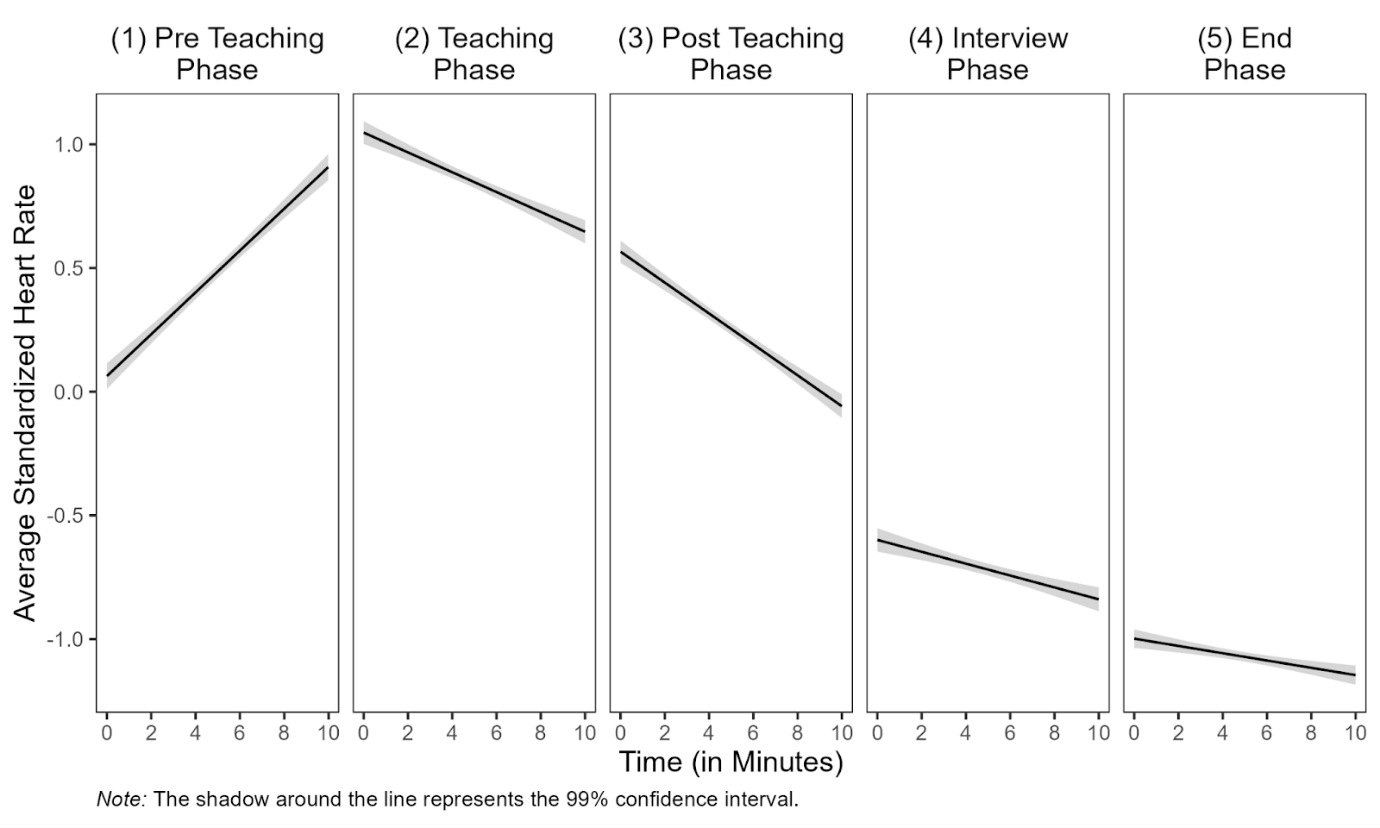
Table XX n, M, SD *of individual intercepts/slopes*

|  |  |  |  |
| --- | --- | --- | --- |
| Phase | n1 | *M* | SD |
| (1) Pre-teaching phase | 6896 | 0.052/0.085\* | 0.820/0.133 |
| (2) Teaching phase | 7150 | 1.025\*/-0.039\* | 0.690/0.108 |
| (3) Post-teaching phase | 6664 | 0.549\*/-0.060\* | 0.547/0.101 |
| (4) Interview phase | 6287 | -0.617\*/-0.022 | 0.614/0.070 |
| (5) End phase | 5990 | -1.004\*/-0.012 | 0.500/0.074 |
| *Note.* \* *p* < .05, 1All measurement time points for all subjects per phase included in the calculations. | | | |

The Table XX shows that mean intercepts differed significantly from zero for all phases except the (1) pre-teaching phase. On average, the mean slopes were negative for all phases but the (1) pre-teaching phase, which means that the subjects’ mean HR increased over the course of the (1) pre-teaching phase. By contrast, the participants’ mean HR decreased in the later course of the study. The mean slope was significantly different from zero for the first three phases.

Fig. XX shows the linear estimation for standardized mean HR for the five different phases. The trend of the entire course of HR can also be represented in this Fig. XX. We see a strong slope in the (1) pre-teaching phase and a decrease in the following phases. In the (4) interview and (5) end phase there is hardly any slope left.

Fig. XX *Linear estimation for standardized mean HR in each phase*



**## Prediction of mean HR with teaching experience and self-report data:**

Correlations between the standardized mean HR and the predictor variables for the different phases are shown in Table XX - XX.

Table XX *Partial correlations between standardized mean HR, teaching experience, disruption factor, confidence factor and gender for (1) pre-teaching phase*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | 2 | 3 | 4 | 5 |
| 1 HR | − .18 | − .10 | − .03 | .18 |
| 2 TE |  | **− .26\*** | **.34\*** | .10 |
| 3 DF |  |  | **− .24\*** | .10 |
| 4 CF |  |  |  | − .07 |
| 5 G |  |  |  | − |
| *Note.* HR = standardized mean heart rate, TE = teaching experience, DF = disruption factor, CF = confidence factor, G = gender. \* *p* < .05, additionally highlighted in bold. | | | | |

Table XX shows that correlations between the standardized mean HR and predictor variables in the (1) pre-teaching phase were only significant between teaching experience and the self-report data (disruption factor: *r* = − .26, confidence factor: *r* = .34). Moreover, the disruption and confidence factor correlated significantly as well (*r* = − .24).

Table XX *Partial correlations between standardized mean HR, teaching experience, disruption factor, confidence factor and gender for (2) teaching phase*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | 2 | 3 | 4 | 5 |
| 1 HR | 0.03 | − .19 | − .02 | .15 |
| 2 TE |  | **− .23\*** | **.36\*** | .07 |
| 3 DF |  |  | **− .25\*** | .11 |
| 4 CF |  |  |  | − .08 |
| 5 G |  |  |  | − |
| *Note.* HR = standardized mean heart rate, TE = teaching experience, DF = disruption factor, CF = confidence factor, G = gender. \* *p* < .05, additionally highlighted in bold. | | | | |

In the (2) teaching phase, the teaching experience correlated significantly with the disruption (*r* = − .23) and the confidence (*r* = − .36) factor and the self-report data with each other (*r* = − .23).

Table XX *Partial correlations between standardized mean HR, teaching experience, disruption factor, confidence factor and gender for (3) post-teaching phase*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | 2 | 3 | 4 | 5 |
| 1 HR | .00 | **.24\*** | .11 | − .10 |
| 2 TE |  | **− .23\*** | **.35\*** | .08 |
| 3 DF |  |  | **− .27\*** | .10 |
| 4 CF |  |  |  | − .07 |
| 5 G |  |  |  | − |
| *Note.* HR = standardized mean heart rate, TE = teaching experience, DF = disruption factor, CF = confidence factor, G = gender. \* *p* < .05, additionally highlighted in bold. | | | | |

Table XX shows that in the (3) post-teaching phase, the dependent variable correlated significantly with the disruption factor (*r* = .24). Additionally, the teaching experience again correlated significantly with the disruption (*r* = − .23) and the confidence (*r* = .35) factor and both factors with each other (*r* = − .27).

Table XX *Partial correlations between standardized mean HR, teaching experience, disruption factor, confidence factor and gender for (4) interview phase*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | 2 | 3 | 4 | 5 |
| 1 HR | .21 | − .05 | − .04 | − .07 |
| 2 TE |  | **− .22\*** | **.36\*** | .09 |
| 3 DF |  |  | **− .25\*** | .08 |
| 4 CF |  |  |  | − .08 |
| 5 G |  |  |  | − |
| *Note.* HR = standardized mean heart rate, TE = teaching experience, DF = disruption factor, CF = confidence factor, G = gender. \* *p* < .05, additionally highlighted in bold. | | | | |

Table XX shows that in the (4) interview phase, the teaching experience again correlated significantly with the disruption (*r* = − .22) and the confidence (*r* = .36) factor and both factors with each other as well (*r* = − .25).

Table XX *Partial correlations between standardized mean HR, teaching experience, disruption factor, confidence factor and gender for (5) end phase*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | 2 | 3 | 4 | 5 |
| 1 HR | .10 | .05 | − .11 | − .18 |
| 2 TE |  | **− .25\*** | **.36\*** | .10 |
| 3 DF |  |  | **− .24\*** | .09 |
| 4 CF |  |  |  | − .10 |
| 5 G |  |  |  | − |
| *Note.* HR = standardized mean heart rate, TE = teaching experience, DF = disruption factor, CF = confidence factor, G = gender. \* *p* < .05, additionally highlighted in bold. | | | | |

In the (5) end phase, the teaching experience again correlated significantly with disruption (*r* = − .25) and the confidence (*r* = .36) factor and both factors with each other as well (*r* = − .24).

The predictions of the subjects’ standardized mean HR for the (2) teaching, the (3) post-teaching, the (4) interview and the (5) end phase with teaching experience are shown in Table XX - XX.

Table XX shows the prediction of HR with teaching experience and self-report data for the (2) teaching phase.

Table XX *Multiple linear regression for (2) teaching phase*.

|  | *Dependent Variable: Standardized mean HR* | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | *(1)* | | *(2)* | | (3) | | (4) | |
|  | Est. | *p* | Est. | *p* | Est. | *p* | Est. | *p* |
| Teaching Experience | 0.005 (0.005) | .34 | 0.002 (0.005) | .73 | 0.005 (0.006) | .42 | 0.003  (0.006) | .67 |
| Disruption Factor |  |  | -0.062 (0.041) | .13 |  |  | -0.065  (0.042) | .13 |
| Confidence Factor |  |  |  |  | 0.004 (0.046) | .92 | -0.014 (0.047) | .76 |
| Constant | **0.813\* (0.057)** | <.01 | **1.150\* (0.227)** | <.01 | **0.778\* (0.349)** | .03 | **1.274\* (0.471)** | .01 |
| *Note*. \* *p* < .05, additionally highlighted in bold, Standard error in parentheses. | | | | | | | | |

The multiple linear regression revealed no significant effects for the (2) teaching phase (see Table XX).

Table XX shows the prediction of HR with teaching experience and self-report data for the (3) post-teaching phase.

Table XX *Multiple linear regression for (3) post-teaching phase*.

|  |  |  | *Dependent Variable: Standardized mean HR* | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | *(1)* | | *(2)* | | (3) | | (4) | |
|  |  |  | Est. | *p* | Est. | *p* | Est. | *p* | Est. | *p* |
|  |  | Teaching Experience | 0.002 (0.005) | .70 | 0.002 (0.005) | .76 | -0.003 (0.006) | .55 | -0.001  (0.006) | .22 |
|  |  | Disruption Factor |  |  | 0.073 (0.040) | .07 |  |  | **0.084\***  **(0.041)** | .04 |
|  |  | Confidence Factor |  |  |  |  | 0.027 (0.045) | .55 | 0.051 (0.046) | .27 |
|  |  | Constant | **0.272\* (0.005)** | <.01 | -0.122 (0.222) | .59 | 0.069 (0.343) | .84 | -0.570 (0.457) | .22 |
|  |  | *Note*. \* *p* < .05, additionally highlighted in bold, Standard error in parentheses. | | | | | | | | |

The multiple linear regression shown in Table XX revealed a significant effect for the (3) post-teaching phase when predicting the HR with teaching experience and controlling for the disruption factor (*b* = 0.084, *p* < .05).

Table XX shows the prediction of HR with teaching experience and self-report data for the (4) interview phase.

Table XX *Multiple linear regression for (4) interview phase*.

|  | *Dependent Variable: Standardized mean HR* | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | *(1)* | | *(2)* | | (3) | | (4) | |
|  | Est. | *p* | Est. | *p* | Est. | *p* | Est. | *p* |
| Teaching Experience | **0.012\* (0.006)** | .03 | 0.011 (0.006) | .06 | **0.013\* (0.006)** | .04 | 0.012  (0.007) | .07 |
| Disruption Factor |  |  | -0.020 (0.044) | .60 |  |  | -0.024  (0.047) | .61 |
| Confidence Factor |  |  |  |  | -0.010 (0.050) | .85 | -0.016 (0.052) | .76 |
| Constant | **0.793\* (0.062)** | <.01 | -0.684 (0.252) | <.01 | -0.721 (0.382) | .06 | -0.541 (0.522) | .30 |
| *Note*. \* *p* < .05, additionally highlighted in bold, Standard error in parentheses. | | | | | | | | |

The multiple linear regression shown in Table XX revealed significant effects for the (4) teaching phase when predicting the HR only with teaching experience (*b* = 0.012, *p* < .05) and teaching experience when controlling for the confidence factor (*b* = 0.013, *p* < .05).

Table XX shows the prediction of HR with teaching experience and self-report data for the (5) end phase.

Table XX *Multiple linear regression for (5) end phase*.

|  | *Dependent Variable: Standardized mean HR* | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | *(1)* | | *(2)* | | (3) | | (4) | |
|  | Est. | *p* | Est. | *p* | Est. | *p* | Est. | *p* |
| Teaching Experience | 0.002 (0.004) | .67 | 0.003 (0.005) | .58 | 0.004 (0.005) | .46 | 0.004 (0.005) | .43 |
| Disruption Factor |  |  | 0.019 (0.035) | .60 |  |  | 0.011 (0.037) | .76 |
| Confidence Factor |  |  |  |  | -0.035 (0.039) | .38 | -0.032 (0.041) | .44 |
| Constant | **-1.075\* (0.049)** | <.01 | **-1.176\* (0.199)** | <.01 | -0.811 (0.300) |  | **-0.897\* (0.411)** | .03 |
| *Note*. \* *p* < .05, additionally highlighted in bold, Standard error in parentheses. | | | | | | | | |

The multiple linear regression revealed no significant effects for the (5) end phase (see Table XX).

Regarding \*\*Hypothesis 2a\*\*, the prediction of the subjects’ mean HR for the (2) teaching, the (3) post-teaching, the (4) interview and the (5) end phase with teaching experience revealed only for the fourth phase, the (4) interview phase, (Table XX, Model 1) a significant result (*b* = 0.012, *p* < .05). When controlling the teaching experience for shared variance with the self-reported data (\*\*Hypothesis 2b\*\*), the disruption factor was a significant predictor for the mean HR in the (3) post-teaching phase (*b* = 0.084, *p* < .05, see Table XX Model 4). When controlling for shared variance with the confidence factor (\*\*Hypothesis 2b\*\*) and considering the three predictors in concert and controlling for their common variance (\*\*Hypothesis 2c\*\*), the models revealed no significant effects.