New approaches to teachers' experience of stress: Do heart rate measurements with fitness trackers provide an efficient, inexpensive, and robust measurement method?

Mandy Klatt[1], Peer Keßler[1], Gregor Kachel[1,2], Christin Lotz[1], & Anne Deiglmayr[1]

[1] Leipzig University

[2] Max Planck Institute for Evolutionary Anthropology

Author Note

Abstract

One or two sentences providing a **basic introduction** to the field, comprehensible to a scientist in any discipline.

Two to three sentences of **more detailed background**, comprehensible to scientists in related disciplines.

One sentence clearly stating the **general problem** being addressed by this particular study.

One sentence summarizing the main result (with the words "**here we show**" or their equivalent).

Two or three sentences explaining what the **main result** reveals in direct comparison to what was thought to be the case previously, or how the main result adds to previous knowledge.

One or two sentences to put the results into a more **general context**.

Two or three sentences to provide a **broader perspective**, readily comprehensible to a scientist in any discipline.

Keywords: heart rate; photoplethysmography; wearable electronic device; teaching

Word count: X

New approaches to teachers' experience of stress: Do heart rate measurements with fitness trackers provide an efficient, inexpensive, and robust measurement method?

Introduction

Physiological data such as heart rate are becoming increasingly important in research on stress experience. They represent an important indicator of physical or emotional stress, as increased workload is associated with increased heart rate (Sachs, 2014). Furthermore, they allow a more objective recording of stress than self-reports (Runge, Haarman, & Fisher, n.d.). However, capturing heart rate in an educational context requires the use of low-cost and non-invasive instruments. Fitness trackers worn on the wrist have the potential to be such a useful tool (Ferguson, Rowlands, Olds, & Maher, 2015).

To date, there is still little evidence on the usefulness of heart rate measurements using fitness trackers in teaching and learning settings (Ertzberger & Martin, 2016; Lowe, 2016). Runge et al. (n.d.) alone examined teacher stress in a relatively small sample ($N = 4$ teachers) and showed that high heart rate, high step count, and little sleep indicate more stress in teachers.

Thus, there remains a lack of robust studies on whether fitness trackers are an efficient, inexpensive, and robust measurement method for assessing teachers' experience of stress. Since, among other things, greater professional experience is associated with lower feelings of stress in the teaching profession (Fisher, 2011), this study will also consider the heart rate measurements of experienced and inexperienced teachers.

Aim of the study

Against this background, we investigated whether heart rate measurements using fitness trackers in a controlled teaching-learning setting are suitable to map differences between (1) anticipation, arousal, and rest phases. We expected heart rates to be higher in

59 the arousal phase than in the rest phases.

Methods

61 We report how we determined our sample size, all data exclusions (if any), all

62 manipulations, and all measures in the study.

Participants

64 The sample consisted in total of $N = 63$ pre- and in-service teachers. The subjects

65 were recruited from the Leipzig University or from German schools in Saxony via personal

66 contact, e-mail lists and flyers. Data of two participants were excluded from the analyses

67 due to insufficient data quality, yielding demographic data from $N = 61$ subjects.

68 The subjects (39 women; 63.93 %) had a mean age of 29.60 years ($SD = 10.40$; range:

69 19-59) and an average teaching experience of 4.66 years ($SD = 9.30$; range: 0-37).

70 18.03% of the subjects were (studying to become) teachers for primary school, 72.13%

71 were (studying to become) teachers for secondary school and 9.84% were (studying to

72 become) teachers for special education needs.

73 All study procedures were carried out in accordance with the ethical standards of the

74 University's Institutional Review Board and the authors received a positive vote on the

75 study procedures from the Ethics Committee Board of Leipzig University. All participants

76 were informed in detail about the aim and intention of the study prior to testing.

77 Participation in the study was voluntary and only took place after written consent has

78 been given.

Material

80 **Teachers' heart rate.** We used a Fitbit Charge 4 to measure the teachers' heart

81 rate. The device was attached 2-finger widths above the ulnar styloid process to the

82   subject's wrist following the manufacturer's instructions. To determine the heart rate, the

83   Fitbit flashes green LEDs many times per second and uses light-sensitive photodiodes to

84   measure the volume changes in the capillaries and then calculates how many times the

85   heart beats per minute (bpm). Data were automatically wireless synced with an iPad via

86   Bluetooth to a Fitbit account, and subsequently, the intraday second-by-second data were

87   exported for each session using the opensource software Pulse Watch (PulseWatch. URL:

88   https://iccir919.github.io/pulseWatch/public/index.html [accessed 2022-08-03]).

89   **Teachers' self-assessment of arousal during the teaching phase.**   The

90   subject's self-assessment of arousal in response to the nine disruptions during the lesson

91   was assessed in a Stimulated Recall Interview that took place after the lesson.

92   For this purpose, two numerical 11-point rating scales were used: (1) The first rating

93   scale collected data of the teacher's subjective perception of disruptiveness of each

94   disruption with the following question: "On a scale of 0-10: how disruptive did you find the

95   event? 0 is not disruptive at all, 10 is extremely disruptive" (2) The second rating scale

96   assessed the teacher's subjective perception of confidence in dealing with the disruptions

97   during the lesson asked as follows: "How confident did you feel in dealing with this event

98   on a scale of 0-10? 0 is not confident at all and 10 is extremely confident"

99   The response format here was purposely chosen to be closed and with several answer

100   options in order to assign numerical values to the assessment and emotional experience of

101   the subjects, and thus to make them measurable and comparable [Eid, Gollwitzer, and

102   Schmitt (2015); doring2016empirische]. The gradations of the rating scale are unambiguous

103   and the intervals between them are identical. The characteristic value was estimated by

104   the subjects immediately after seeing the recording of the respective disruption and

105   communicated verbally.

Procedure

The data collection was part of a larger research project with a planned sample of n = 40 in-service teachers and n = 40 pre-service teachers.

In a controlled laboratory setting, heart rate data in beats per minute (bpm) were recorded using Fitbit Charge 4 over a total period of approximately two hours. Within this time frame, teachers taught a 15-minute self-prepared lesson to an audience of three actors. For analyzing the heart rate data, we selected five 10-minute intervals of theoretical interest: (1) the pre-teaching phase, which started from the moment the subject put on the watch (2) the teaching phase, which began at the moment when the subject started the lesson (3) the post-teaching phase, which was the 10 minutes immediately after the end of the lesson (4) the interview phase, which was 10 minutes in the middle of the interview (for this, the difference was calculated 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 and 5 minutes were subtracted to get to the start of the 10-minute interval), (5) the end phase, which was the 10 minutes before the subject took off the watch.

In a laboratory setting, three trained actors performed teaching disruptions in a counter balanced fashion while the subject taught a 15-minute micro-teaching unit prepared in advance. To record the subject's heart rate, the Fitbit Charge 4 was put on at least 10 minutes before the start of the lesson. The lesson was recorded by four cameras and an audio recorder. In addition, the subject wore eye-tracking glasses to record gaze behavior.

Subsequently, the subjects 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 using a validated questionnaire (EMU, Helmke et al. (2014)) and self developed scales on the teacher's presence behavior derived from the research literature via the online survey website SoSci Survey (4-point Likert scale; 1 = strongly disagree; 2 = disagree; 3 = agree; 4 = strongly agree). The completion of the

132 questionnaire took approximately 5 minutes.

133 The experimenter then conducted a Stimulated Recall Interview (SRI), where the

134 subject commented and rated the reactions to classroom events while watching the eye

135 tracking video.

136 Finally, a Situational Judgment Test (SJT, Gold and Holodynski (2015)) was used to

137 assess the subject's strategic knowledge of classroom management. The subject was asked

138 to judge alternative actions on school scenarios on a 6-point rating scale from 1 (A) to 6

139 (F) according to school grades. Data from the SJT were again collected as an online

140 questionnaire via the website www.soscisurvey.de and lastet approximately 10 minutes.

141 The Fitbit watch was removed only after the last questionnaire to obtain heart rate

142 data during the entire study.

143 **Data analysis**

144 We used R (Version 4.1.3; R Core Team, 2022) and the R-packages *broom* (Version

145 1.0.1; Robinson, Hayes, & Couch, 2022), *cowplot* (Version 1.1.1; Wilke, 2020), *DescTools*

146 (Version 0.99.45; Andri et mult. al., 2022), *dplyr* (Version 1.0.10; Wickham, François,

147 Henry, & Müller, 2022), *forcats* (Version 0.5.1; Wickham, 2021), *ggplot2* (Version 3.3.5;

148 Wickham, 2016), *ggpubr* (Version 0.4.0; Kassambara, 2020), *ggthemes* (Version 4.2.4;

149 Arnold, 2021), *gridExtra* (Version 2.3; Auguie, 2017), *imputeTS* (Version 3.2; Moritz &

150 Bartz-Beielstein, 2017), *janitor* (Version 2.1.0; Firke, 2021), *jtools* (Version 2.2.0; Long,

151 2022), *lm.beta* (Version 1.6.2; Behrendt, 2022), *lme4* (Version 1.1.30; Bates, Mächler,

152 Bolker, & Walker, 2015), *ltm* (Version 1.2.0; Rizopoulos, 2006), *lubridate* (Version 1.8.0;

153 Grolemund & Wickham, 2011), *MASS* (Version 7.3.55; Venables & Ripley, 2002), *Matrix*

154 (Version 1.5.1; Bates, Maechler, & Jagan, 2022), *msm* (Version 1.6.9; Jackson, 2011), *needs*

155 (Version 0.0.3; Katz, 2016), *papaja* (Version 0.1.0.9999; Aust & Barth, 2020), *polycor*

156 (Version 0.8.1; Fox, 2022), *ppcor* (Version 1.1; Kim, 2015), *purrr* (Version 0.3.4; Henry &

₁₅₇ Wickham, 2020), *readr* (Version 2.1.2; Wickham, Hester, & Bryan, 2022), *readxl* (Version

₁₅₈ 1.4.0; Wickham & Bryan, 2022), *rstatix* (Version 0.7.0; Kassambara, 2021), *stringr*

₁₅₉ (Version 1.4.0; Wickham, 2019), *tibble* (Version 3.1.6; Müller & Wickham, 2021), *tidyr*

₁₆₀ (Version 1.2.0; Wickham & Girlich, 2022), *tidyverse* (Version 1.3.1; Wickham et al., 2019),

₁₆₁ *tinylabels* (Version 0.2.3; Barth, 2022), *viridis* (Version 0.6.2; Garnier et al., 2021a, 2021b),

₁₆₂ and *viridisLite* (Version 0.4.0; Garnier et al., 2021b) for all our analyses.

₁₆₃ **FitBit.** All participants were given a FitBit Smart Watch Charge 4 to wear during

₁₆₄ the experiment.

₁₆₅ **Heart Rate**

₁₆₆ **Results**

₁₆₇ **Discussion and implication**

References

168

169 Andri et mult. al., S. (2022). *DescTools: Tools for descriptive statistics.* Retrieved from

170   https://cran.r-project.org/package=DescTools

171 Arnold, J. B. (2021). *Ggthemes: Extra themes, scales and geoms for 'ggplot2'.* Retrieved

172   from https://CRAN.R-project.org/package=ggthemes

173 Auguie, B. (2017). *gridExtra: Miscellaneous functions for "grid" graphics.* Retrieved from

174   https://CRAN.R-project.org/package=gridExtra

175 Aust, F., & Barth, M. (2020). *papaja: Prepare reproducible APA journal articles with R*

176   *Markdown.* Retrieved from https://github.com/crsh/papaja

177 Barth, M. (2022). *tinylabels: Lightweight variable labels.* Retrieved from

178   https://cran.r-project.org/package=tinylabels

179 Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects

180   models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

181   https://doi.org/10.18637/jss.v067.i01

182 Bates, D., Maechler, M., & Jagan, M. (2022). *Matrix: Sparse and dense matrix classes and*

183   *methods.* Retrieved from https://CRAN.R-project.org/package=Matrix

184 Behrendt, S. (2022). *Lm.beta: Add standardized regression coefficients to lm-objects.*

185   Retrieved from https://CRAN.R-project.org/package=lm.beta

186 Eid, M., Gollwitzer, M., & Schmitt, M. (2015). *Statistik und forschungsmethoden: Mit*

187   *online-materialien (4., überarb. Und erw. aufl.).* Weinheim: Beltz. Retrieved from

188   https://contentselect. com/media/moz_viewer . . . .

189 Ertzberger, J., & Martin, F. (2016). Use of fitness bands by teachers in the classroom.

190   *TechTrends*, *60*(4), 392–397.

191 Ferguson, T., Rowlands, A. V., Olds, T., & Maher, C. (2015). The validity of

192   consumer-level, activity monitors in healthy adults worn in free-living conditions: A

193   cross-sectional study. *International Journal of Behavioral Nutrition and Physical*

194   *Activity*, *12*(1), 1–9.

195   Firke, S. (2021). *Janitor: Simple tools for examining and cleaning dirty data.* Retrieved
196   from https://CRAN.R-project.org/package=janitor

197   Fisher, M. H. (2011). Factors influencing stress, burnout, and retention of secondary
198   teachers. *Current Issues in Education, 14*(1).

199   Fox, J. (2022). *Polycor: Polychoric and polyserial correlations.* Retrieved from
200   https://CRAN.R-project.org/package=polycor

201   Garnier, Simon, Ross, Noam, Rudis, Robert, ... Cédric. (2021a). *viridis -*
202   *colorblind-friendly color maps for r.* https://doi.org/10.5281/zenodo.4679424

203   Garnier, Simon, Ross, Noam, Rudis, Robert, ... Cédric. (2021b). *viridis -*
204   *colorblind-friendly color maps for r.* https://doi.org/10.5281/zenodo.4679424

205   Gold, B., & Holodynski, M. (2015). Development and construct validation of a situational
206   judgment test of strategic knowledge of classroom management in elementary schools.
207   *Educational Assessment, 20*(3), 226–248.

208   Grolemund, G., & Wickham, H. (2011). Dates and times made easy with lubridate.
209   *Journal of Statistical Software, 40*(3), 1–25. Retrieved from
210   https://www.jstatsoft.org/v40/i03/

211   Helmke, A., Helmke, T., Lenske, G., Pham, G., Praetorius, A.-K., Schrader, F.-W., &
212   AdeThurow, M. (2014). Unterrichtsdiagnostik mit EMU. *Aus-Und Fortbildung Der*
213   *Lehrkräfte in Hinblick Auf Verbesserung Der Diagnosefähigkeit, Umgang Mit*
214   *Heterogenität Und Individuelle Förderung*, 149–163.

215   Henry, L., & Wickham, H. (2020). *Purrr: Functional programming tools.* Retrieved from
216   https://CRAN.R-project.org/package=purrr

217   Jackson, C. H. (2011). Multi-state models for panel data: The msm package for R. *Journal*
218   *of Statistical Software, 38*(8), 1–29. Retrieved from https://www.jstatsoft.org/v38/i08/

219   Kassambara, A. (2020). *Ggpubr: 'ggplot2' based publication ready plots.* Retrieved from
220   https://CRAN.R-project.org/package=ggpubr

221   Kassambara, A. (2021). *Rstatix: Pipe-friendly framework for basic statistical tests.*

Retrieved from https://CRAN.R-project.org/package=rstatix

Katz, J. (2016). *Needs: Attaches and installs packages.* Retrieved from
https://CRAN.R-project.org/package=needs

Kim, S. (2015). *Ppcor: Partial and semi-partial (part) correlation.* Retrieved from
https://CRAN.R-project.org/package=ppcor

Long, J. A. (2022). *Jtools: Analysis and presentation of social scientific data.* Retrieved
from https://cran.r-project.org/package=jtools

Lowe, H. (2016). Educational behavior apps and wearable devices: Current research and
prospects. *Educational Technology*, 30–33.

Moritz, S., & Bartz-Beielstein, T. (2017). imputeTS: Time Series Missing Value
Imputation in R. *The R Journal*, *9*(1), 207–218. https://doi.org/10.32614/RJ-2017-009

Müller, K., & Wickham, H. (2021). *Tibble: Simple data frames.* Retrieved from
https://CRAN.R-project.org/package=tibble

R Core Team. (2022). *R: A language and environment for statistical computing.* Vienna,
Austria: R Foundation for Statistical Computing. Retrieved from
https://www.R-project.org/

Rizopoulos, D. (2006). Ltm: An r package for latent variable modelling and item response
theory analyses. *Journal of Statistical Software*, *17*(5), 1–25. Retrieved from
https://doi.org/10.18637/jss.v017.i05

Robinson, D., Hayes, A., & Couch, S. (2022). *Broom: Convert statistical objects into tidy
tibbles.* Retrieved from https://CRAN.R-project.org/package=broom

Runge, N., Haarman, S., & Fisher, M. H. (n.d.). *Using fitbit fitness trackers to measure
teacher stress and coping.*

Sachs, S. (2014). *Physiologische parameter zur bewertung der lernwirksamkeit von
lernsituationen* (PhD thesis). Universität Ulm.

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with s* (Fourth). New
York: Springer. Retrieved from https://www.stats.ox.ac.uk/pub/MASS4/

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis.* Springer-Verlag New York. Retrieved from https://ggplot2.tidyverse.org

Wickham, H. (2019). *Stringr: Simple, consistent wrappers for common string operations.* Retrieved from https://CRAN.R-project.org/package=stringr

Wickham, H. (2021). *Forcats: Tools for working with categorical variables (factors).* Retrieved from https://CRAN.R-project.org/package=forcats

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., ... Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wickham, H., & Bryan, J. (2022). *Readxl: Read excel files.* Retrieved from https://CRAN.R-project.org/package=readxl

Wickham, H., François, R., Henry, L., & Müller, K. (2022). *Dplyr: A grammar of data manipulation.* Retrieved from https://CRAN.R-project.org/package=dplyr

Wickham, H., & Girlich, M. (2022). *Tidyr: Tidy messy data.* Retrieved from https://CRAN.R-project.org/package=tidyr

Wickham, H., Hester, J., & Bryan, J. (2022). *Readr: Read rectangular text data.* Retrieved from https://CRAN.R-project.org/package=readr

Wilke, C. O. (2020). *Cowplot: Streamlined plot theme and plot annotations for 'ggplot2'.* Retrieved from https://CRAN.R-project.org/package=cowplot