Received: 11 December 2022 | Accepted: 12 December 2022

DOI: 10.1111/bjep.12575

the british psychological society
promoting excellence in psychology

ARTICLE

Mind and body in students' and teachers' engagement: New evidence, challenges, and guidelines for future research

Reinhard Pekrun[1,2,3]

[1]Department of Psychology, University of Essex, Colchester, UK

[2]Institute for Positive Psychology and Education, Australian Catholic University, Sydney, Australia

[3]Department of Psychology, University of Munich, Munich, Germany

Correspondence
Reinhard Pekrun, Department of Psychology, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK.
Email: pekrun@lmu.de

Abstract

Background and Aims: Traditionally, research in educational psychology has neglected the physiological foundations of motivation, emotion, engagement, and learning. Recent studies have made substantial progress to more fully consider physiological processes, as documented in the contributions to this special issue. In this commentary, I summarize their findings, discuss strengths and weaknesses, and outline directions for future research.

Results: The studies showcase how physiological indicators can be integrated in research in educational psychology. The resulting findings document links between cardiovascular, electrodermal, and hormonal parameters as well as physical activity and a range of mental and behavioural processes in educational settings. Together, they attest to the critical role of physiological processes in students' and teachers' engagement. However, most of the studies used small samples and correlational designs, and not all of the findings were consistent.

Future Directions: To inform theory and practice in evidence-based ways, we need to make further headway in building a cumulative, coherent knowledge base. To this end, it may be helpful to more precisely specify the status of physiological indicators; secure construct symmetry of physiological, mental, and behavioural variables; use causal designs and within-person analysis; include sufficiently powered samples of participants and measurement occasions; employ multiple indicators and assessments to increase reliability and specificity; define the time windows and lags of assessments that are suited to capture physiological processes and their functions; and consider the role of socio-cultural contexts.

KEYWORDS
academic achievement, cortisol, educational psychology, electrodermal activity, emotion, engagement, heart rate, motivation, physical activity, review

INTRODUCTION

As aptly argued by Martin et al. (2022/this issue), scientists in educational psychology have considered motivation, emotion, engagement, and learning as mental phenomena but have neglected their physiological foundations. Physiological variables have been neglected despite the fact that they are core constituents of major constructs in the field. A case in point is emotions, which are typically defined as multi-component processes that comprise not only mental but also physiological components (e.g., Scherer & Moors, 2019). Consistent with this neglect, researchers have focused on using self-report to assess motivation, emotion, and engagement, but have failed to integrate physiological measures more systematically in their studies. Certainly, self-report has several clear advantages (Pekrun, 2020). Administering self-report measures is less costly than alternative methods, and it is suited to capture nuances of subjective experiences that other methods cannot assess. Nevertheless, an exclusive use of self-report limits the range of possible conclusions. Self-report is restricted to assess processes that are accessible and can be consciously represented, and it is subject to memory biases and response sets. As such, from a methodological perspective as well, it is imperative to complement mentally focused measures with indicators that represent bodily responses. Such indicators can not only make unconscious processes accessible; they typically also are less obtrusive and less subject to tendencies to respond in socially desirable ways.

By including measures of bodily processes, the 12 studies reported in this special issue represent significant advances over the traditional state of the art. The studies considered different measures of arousal, stress, and physical activity, and related the resulting scores to a range of constructs of motivation, emotion, engagement, and learning, in both students and teachers. Collectively, these studies showcase how physiological indicators can be combined with other measures, including self-report, behavioural observation, and performance scores. In the following, I will first summarize major findings of the 12 studies, organized by the physiological variables assessed. I will outline strengths of these studies, but will also address limitations that are typical for the field. I will then discuss how these limitations can be used to develop guidelines for future research that may be suited to further advance the field and launch a cumulative science of physiological processes in educational settings.

OVERVIEW OF MAJOR FINDINGS

Cardiovascular indexes of arousal and regulation

Four of the studies presented in this special issue used cardiovascular parameters including heart rate (HR), heart rate variability (HRV), and cardiac vagal tone (CVT) in relation to measures of emotions, engagement and achievement. HR is commonly interpreted as a measure of physiological arousal. At least in some contexts, HRV and CVT can be interpreted as general indicators of ability to self-regulate.

Ketonen et al. (2022/this issue) used within-person analysis to examine relations between HR and HRV, on the one hand, and emotions differing in valence and arousal, on the other, in a sample of upper secondary school students. Emotions were assessed using experience sampling methodology (ESM) over a period of 10 days. The findings show that excitement was positively related to HR and negatively related to HRV, whereas calmness and boredom were negatively related to HR and positively related to HRV. Although these correlations were small, they document consistent links between emotions and cardiovascular indexes of arousal. Mainhard et al. (2022/this issue) employed HR to operationalize secondary school teachers' effort in managing closeness to students, defined as the within-person correlation of HR and observed closeness. In subsequent between-person analysis, this measure of teachers' effort, which can be interpreted as an index of emotional labour, was negatively related to students' perceptions of teacher closeness. In addition, effort related positively to students' self-reported anger and anxiety and negatively to their relaxation; these latter correlations were not significant but consistent in direction. Overall, the findings demonstrate that HR can be used to develop unobtrusive measures of emotional labour.

Zaccoletti et al. (2022/this issue) used an experimental manipulation to investigate the effects of worry on text comprehension in a sample of 7th graders. Memory updating ability and resting state HRV were used as non-experimental predictors. In linear mixed modelling, HRV had a negative main effect on performance. In addition, HRV interacted with the experimental conditions; HRV had a negative effect on performance in the worry condition, but a positive effect in the non-worry condition. For understanding the findings, it is important to know that participants may have interpreted the worry induction as a challenge rather than a threat; performance was higher in the worry condition than in the non-worry condition. In a two-wave longitudinal study over one school year, Mastromatteo et al. (2022/this issue) examined relations between elementary school students' cardiac vagal tone (CVT; assessed via HRV) at rest and during a demanding cognitive task, their engagement, and their perceptions of classroom climate. CVT showed a developmental increase over the year. Furthermore, the rate of change was positively related to students' self-reported engagement, and at the first wave, the difference between task-related and baseline CVT (i.e., an adaptive CVT response to the task) was linked to better engagement when climate perceptions were negative. These findings document that CVT indexes can be linked in meaningful ways to self-report measures of engagement.

Electrodermal activity (EDA) indexes of arousal

Two studies used EDA parameters as indicators of arousal. Roos et al. (2022/this issue) assessed 8th grade students' perceptions of control, frequency of skin conductance responses (SCRs), self-reported anxiety components, and performance during six blocks of a low-stakes math test. Within-person analysis showed that control was negatively related to anxiety, including its physiological component, but positively to SCR frequency. Post-hoc analyses showed trends for control being negatively related to SCRs when anxiety was the dominant emotion, but positively related to SCRs when joy was dominant. This finding suggests that EDA can be used as an indicator of the arousal component of specific emotions if combined with other (emotion-specific) measures.

Törmänen et al. (2022/this issue) used SCR frequency to assess the arousal dimension of affect in collaborative groups of secondary school students working on science problems. Group interaction was observed and coded for socio-emotional interaction episodes, valence of group affect (based on verbal and bodily expression), and regulation of emotions. Clusters of episodes defined by valence and regulation differed in patterns of arousal. Arousal showed more variation in clusters of episodes with positive group affect or frequent regulation, possibly due to variations in mental effort during these episodes as reflected in fluctuations in SCRs.

Hormones and brain activation as biomarkers of stress

In a two-wave study with 7th and 8th grade students, Hoferichter and Raufelder (2022/this issue) assessed cortisol, alpha amylase, and oxidative stress on one day during the first half of the school year, and self-reported grit as well as school engagement both during the first and the second half of the year. Most of the synchronous and lagged relations among the biomarkers and with grit and engagement were not significant. Two exceptions were a negative effect of cortisol on Time 2 cognitive engagement, and a negative effect of alpha amylase on Time 2 grit. Graham et al. (2022/this issue) assessed cortisol level and task-related expectancy of success before, during, and after a midterm practice examination taken by engineering students. As hypothesized, cortisol levels decreased after the start of the examination. The correlations of cortisol levels over time were low, and the concurrent correlations between expectancy and cortisol levels were near zero. However, latent growth modelling showed that students who were better able to keep up positive expectancies during the examination showed a steeper decline in cortisol, suggesting that expectancy protected them against suffering from stress as indicated by insufficient recovery from initial levels of cortisol.

In an intervention study with 5th and 6th graders, Dettweiler et al. (2022/this issue) compared outdoor teaching (one day per week in the forest) with a control group taught indoors. Outdoor teaching provided exposure to nature but also more interdisciplinarity of the curriculum, physical activity, and student autonomy. Participants in the intervention group showed lower end-of-morning levels of cortisol, attesting to beneficial effects of the intervention. In addition, relations of group membership with maturation of limbic brain areas and activation of these areas under stress emerged. Interpretation was limited by age differences between the two groups, the small size especially of the control group ($n = 11$), and extreme cases in this group. Finally, Jõgi et al. (2022/this issue) assessed first grade classroom teachers' levels of cortisol in the middle of the school morning, slope of cortisol levels from awakening to bedtime, self-reported chronic levels of stress related to teaching, and child- versus teacher-centred classroom instruction as well as their students' math achievement. Cortisol was assessed on two working days, and instruction was observed during three lessons. Self-reported stress was negatively related to child-centred teaching. Counter to expectations, neither stress nor teaching style were related to students' achievement. Furthermore, cortisol level and slope did not show significant relations with perceived stress, teaching practices, or student achievement.

Physical activity

De Bruijn et al. (2022/this issue) examined physical activity and motivation during one physical education lesson in a sample of 3rd and 4th grade students. Physical activity was assessed with accelerometers, motivation with a combined scale measuring perceived value, interest, competence, and effort during the lesson. Motivation was positively related to light and moderate activity. Possible explanations not only include effects of motivation on behaviour but also reverse effects of activity on motivation. Students' overall academic achievement was positively related to their motivation during the lesson, but did not moderate the relation between motivation and activity as expected by the authors.

Heemskerk et al. (2022/this issue) used multilevel modelling to investigate relations of physical activity during physical education lessons with self-reported affect and observed task behaviour during subsequent lessons in a sample of 3rd to 5th graders. Light physical activity negatively predicted subsequent on-task behaviour and positively predicted active off-task behaviour. Moderate activity positively predicted on-task behaviour and negatively predicted passive off-task behaviour. Most of the relations of physical activity with affect and tiredness during the subsequent lesson were not significant. Overall, the findings suggest beneficial effects of moderate activity, although the presumed mediational pathway through affect could not be documented.

Conclusions

Taken together, two important messages follow from the findings of the 12 studies. First, each of the studies considered a specific combination of physiological, mental, and behavioural variables. As such, it is not possible to draw more general multi-study conclusions about relations between single variables—such as the link between arousal and anxiety, or the relation between physical activity and achievement—from the findings. However, as a set, the studies clearly document that physiological variables can be related in meaningful ways to mental variables, behaviour, and performance. They show how emotions and engagement relate to physiological arousal; how different educational contexts, such as examinations and outdoor experiences, can influence physiological functioning; and how students' physical activity can relate to their academic motivation and engagement.

Second, it is evident that research on physiological processes has the potential to inform educational practice. For example, knowledge about the impact of examinations on the release of stress hormones could guide the design of assessments; knowledge about the effects of physical activity on students' mental effort and on-task behaviour could guide the design of physical activity components in the curriculum; and knowledge about the impact of emotions in educational contexts on immune system processes could motivate efforts to design educational institutions and learning environments in healthier ways.

In sum, the studies convincingly demonstrate how physiological indicators can be integrated in scientific inquiry in educational psychology, and they yielded a wealth of exciting findings. Together these findings document that it is imperative to consider physiological processes to more fully understand students' and teachers' motivation, emotions, and engagement. However, we also need to consider that this is an emerging field of research. Multiple replications and cumulative evidence will be needed to make sure that the results are generalizable and can be used to derive firm recommendations for practice. The findings of the current studies are promising, but they also suggest that this is not a trivial task. Theoretically expected relations did not always show up, and the observed relations were often weak and sometimes not consistent. To understand the reasons, it may be helpful to interpret the size and consistency of observed relations relative to critical issues in theory, study design, and measurement. I will turn to these issues next, and I will also discuss how they can be used to develop guidelines for future research.

CONCEPTUAL FOUNDATIONS

Relations between mind, body, and behaviour

Should we consider physiological processes as indicators, components, predictors, origins or outcomes of motivation, emotions, and engagement? This question needs to be answered to meaningfully design studies including physiological parameters and to interpret their findings. Answers may differ according to the constructs considered. For example, the relation between physiological processes and emotions can be conceptualized in at least three different ways. First, following contemporary multi-component definitions of emotion, physiological activation or deactivation can be considered as a component of emotions (e.g., as a component of test anxiety in the study by Roos et al., 2022/this issue). Second, physiological processes can be considered as antecedents of emotions, as in Heemskerk et al.'s (2022/this issue) study investigating effects of physical activity on affective states (presumably, these effects are at least in part due to the physiological processes prompted by physical activity). Third, physiological processes can be triggered by emotions; an example is the effects of emotions on immune system functioning and health (Pekrun et al., 2023).

In all three cases, it is important to attend to the multi-component nature of emotions. Emotions are best considered as systems of coordinated, but only loosely coupled responses to important events (e.g., Moors, 2017). By implication, correlations between physiological parameters and non-physiological components of emotions, or with wholistic overall measures of emotions, cannot be expected to be strong. The findings reported by Roos et al. (2022/this issue) are in line with this view—the links between physiological arousal and other components of test anxiety were consistent, but relatively small. Furthermore, effects of antecedent physiological processes on subsequent emotions may be best detected when specifically considering effects on the physiological component of emotions, as relations with overall measures of emotions may mask these effects. Similarly, when investigating effects of emotions on physiological processes, it may be best to differentiate between the effects of different components of emotions.

Similar principles hold for related constructs, such as motivation and engagement. Typically, constructs of motivation and engagement are multidimensional. For targeting the role of physiological processes, it is important to specify these multiple dimensions and consider if they are linked in different ways to bodily processes.

Principles of construct symmetry

Following Brunswik (1955) and Ajzen (1988; Ajzen & Fishbein, 2005), it is important to attend to the conceptual symmetry of constructs when investigating their relations. Constructs can be specific in terms of denoting a single psychological process in a single situation at a given point in time. An example is a student's thoughts about possible failure during a math examination on a given day. Alternatively, constructs can be generalized across processes, situations, and time. An example is trait neuroticism, which denotes people's general disposition to process various kinds of negative thoughts and emotions in different situations and across time points. Other constructs are located in between these extremes by showing a medium level of generality, such as students' math anxiety which pertains to one specific emotion (anxiety) in one class of situations (dealing with math), but can be generalized over time.

Theory and empirical evidence suggest that relations between constructs are stronger when the constructs are symmetrical, that is, when they show the same level of specificity (or generality; e.g., Kretzschmar et al., 2018). This principle of construct symmetry has profound implications for investigating links between physiological, mental, and behavioural processes. For example, if we want to examine relations between physiological arousal and cognitive performance, then chances to detect these relations may be low if we assess arousal using one single parameter (e.g., SCR frequency) during one single experimental session including one specific type of task, but performance in terms of students' GPA over the entire school year and across a broad range of academic subjects. Chances are much better if we link the observed SCRs to performance on the task itself, or if we observe arousal using several indicators over an extended period of time and then link it to students' overall performance.

A few of the studies reported in this special issue attended to construct symmetry. For example, the studies that assessed physiological indicators and mental or behavioural processes in parallel over time at least made sure that the constructs were matched in terms of temporal specificity (Graham et al., 2022/this issue; Roos et al., 2022/this issue). However, in the majority of the studies, the target constructs did not show a strong match. Many of the studies assessed one single physiological indicator in a single session (or on a limited number of days), but related the scores to broad constructs of engagement and achievement. The resulting lack of construct symmetry needs to be considered when interpreting the findings. Some of the null results reported in the studies may have become significant when assessing physiological markers across a broader, more representative range of situations. Alternatively, for single physiological markers assessed in a specific situation, relations may have been stronger when considering similarly specific measures of engagement.

Issues of construct symmetry pose a conundrum for psychophysiological and neuroscientific research given the costs of assessing physiological indicators. This may be less of an issue when confining both physiological and mental or behavioural assessment to a specific task. It is a problem when the goal is to predict students' and teachers' overall engagement, achievement, and well-being, across situations and time.

EMPIRICAL PARADIGMS

Correlational and causal study designs

In early stages of research in a given field, it is important to first establish that variables are related, before proceeding to examine the causal processes that generate the relations. Following this logic, most of the studies reported in this special issue used correlational designs in which physiological processes and other variables were either assessed synchronously, or in a lagged fashion but without controlling for autoregressive effects and confounders (exceptions are the longitudinal analysis by Hoferichter & Raufelder, 2022/this issue, and the experimental studies by Dettweiler et al., 2022/this issue, and Zaccoletti et al., 2022/this issue). Correlational analysis is an important step forward which lays the foundations for subsequent research. However, from correlational designs, we cannot infer any conclusions about the causal status of physiological variables. This is not only true for designs that investigate associations between variables at a given point in time, but also for longitudinal data on change that are used to examine the association between parallel change in different variables. For example, the data reported by Graham et al. (2022/this issue) inform us how change in expectancy of success and change in cortisol levels were associated across an examination, but they do not inform us about the temporal or causal ordering of these changes. As such, most of the findings reported in the current studies remain open to interpretation as to the direction and causal nature of the relations between physiological and other variables.

To make further headway in this field of research, it will be important to move beyond correlational studies and use causal designs, including experiments as well as controlled longitudinal studies. This is important from both theoretical and practical perspectives. From correlational data, it is difficult to develop sound theories of change. Using data on mere associations between variables, it would be problematic and potentially misleading to inform practitioners how they should change their practices, and what the effects of change might be.

Between- and within-person research

Physiological processes occur within persons, and theories on their functions for motivation, emotions, and engagement typically refer to within-person mechanisms. Even when the goal is to explain individual differences (e.g., differences between students in academic achievement), the presumed pathways generating these differences are located within persons. As such, it would be straightforward to use within-person analytic designs to investigate the role of physiological processes for motivation and engagement. Nevertheless, studies in the field typically use between-person designs. Again, the costs incurred by assessing physiological indicators multiple times in sufficiently large samples of participants may be a major reason. However, similar to issues of construct asymmetry, using between-person designs poses a problem: From between-person findings, we cannot infer how individuals function, given that between- and within-person relations between variables are statistically independent and can diverge widely (Murayama et al., 2017; Orth et al., 2021; Voelkle et al., 2014).

Most of the studies documented in this special issue focused on between-person analysis to infer within-person functioning, using either correlational or experimental between-person designs. Notable exceptions are the within-person analyses reported by Heemskerk et al. (2022/this issue), Ketonen et al. (2022/this issue), and Roos et al. (2022/this issue). To make further headway in the field, it will be important to follow the model set by these studies and more broadly complement between-person analysis by within-person analysis. This will make it not only possible to directly examine within-person mechanisms, but also to investigate their generalizability across persons. In the past years, both fixed-effect and random-effects models have been developed that can be used. For example, the random-intercept cross-lagged panel model (RI-CLPM) proposed by Hamaker et al. (2015) allows investigation of fixed within-person effects. Dynamic structural equation modelling (DSEM) goes one step further by estimating both fixed and random effects (Hamaker et al., 2018). The RI-CLPM has the clear advantage that it can be used with a small number of assessments (three waves can be sufficient), whereas random-effect approaches like DSEM require more assessments within persons (Schultzberg & Muthén, 2018).

All types of physiological research that use multiple or continuous assessments of indicators over time may be ideally suited to use modelling approaches that require large numbers of assessments. However, to this end, the indicators of mental processes or behaviour also need to be assessed multiple times. Furthermore, to detect lagged effects linking different variables, the time intervals used for the analysis need to be properly calibrated. If the intervals are too short, there may not be sufficient change, which can result in large autoregressive effects leaving little residual variance to be explained by other variables (see van Vemde et al., 2022, for high autocorrelations in second-to-second observation of teacher and student behaviour). Aggregating scores across intervals may be a solution.

By assessing both between- and within-person relations, and both fixed and random within-person relations, approaches like DSEM allow integrating idiographic and nomothetic perspectives on physiology, motivation, and engagement. The resulting findings may pose challenges for the field. In our theories, we typically assume universality of basic psychological mechanisms, including their links with physiological foundations (e.g., Pekrun, 2006; Vansteenkiste et al., 2020). For many important relations, there is sound evidence for universality across genders, academic domains, and socio-cultural contexts (e.g., the relations between students' anxiety and their achievement; see Barroso et al., 2021; Pekrun, 2009). However, this type of universality refers to generalizability of sample statistics across groups of persons and contexts. It remains largely unknown to what extent these relations are generalizable across single, individual persons. The tools that are available today should be used to find answers to this fundamentally important question.

INCREASING THE QUALITY OF EMPIRICAL ASSESSMENTS

Sample size: Participants and measurement occasions

Given the costs, it is often difficult to conduct physiological assessments with large numbers of participants within single studies. This was also true for the studies documented in this special issue. With one exception (the physical activity study by de Bruijn et al., 2022/this issue, N = 891), all of the studies used small participant samples (Ns < 100). This may be less of a problem for studies that use multilevel analysis based on multiple assessments per person. For multilevel analysis, power is jointly determined by the size of the participant sample and the size of the sample of measurement occasions (e.g., Schultzberg & Muthén, 2018). Especially for studies that primarily rely on the size of participant samples, however, small size of these samples can be a serious problem.

With small sample size, the risk is high that the observed findings are due to chance, and they show higher variability (especially for small effects). Estimates based on small samples can be inflated, and the likelihood of replication is reduced (e.g., Ioannidis, 2005; Schönbrodt & Perugini, 2013; see also Fiedler, 2011). Given these problems, the field needs to move forward by securing sufficient sampling, preferably not only for participants but also for indicators, settings, and measurement occasions. One way to achieve this goal is to secure funding for large-scale physiological studies in education. Alternatively, we could use the multi-lab and replication strategies that other fields, such as social and developmental psychology, have developed to increase the power of empirical investigation (see, e.g., Byers-Heinlein et al., 2021; Nosek & Lakens, 2014).

Reliability, sensitivity, and specificity

Physiological indicators differ in psychometric quality. To properly estimate links with mental processes and behaviour, they need to be sufficiently reliable, sensitive, and specific (of course, the same is true for mental and behavioural measures). For the psychophysiological indicators typically used in the field, especially reliability and specificity are potential problems. Reliability can be a problem when indicators show random fluctuation over time. If this is the case, then any single measurement may represent random variance rather than systematic variance, and true relations with other variables may be underestimated. Lack of reliability in terms of low retest stability is frequently observed for physiological parameters, especially if the time interval extends beyond a few minutes. An example is the low correlations of cortisol level over time reported by Graham et al. (2022/this issue). However, there also are exceptions, such as the relatively high retest stability of CVT from baseline to task completion reported by Mastromatteo et al. (2022/this issue; for low reliability of HRV over longer time intervals, see Weiner & McGrath, 2017). Lack of specificity is typical for most biomarkers of arousal and stress—parameters like HR, HRV or

SCR frequency may be sensitive to various mental processes, including emotions, cognitive challenge, and mental effort.

Depending on physiological process and context, reliability may be secured by increasing the number of measurements (similar to increasing the reliability of an achievement test by including multiple items). To establish specificity, it may be necessary to combine the respective physiological indicator with other measures. An example cited earlier is the study by Roos et al. (2022/this issue); in this study, SCR frequency could be interpreted as an indicator of anxiety by combining it with self-reports of specific emotions.

Time intervals and time lags

Continuous assessments of physiological indicators need to be aggregated over time to examine their relations with mental or behavioural variables. How to best choose appropriate time intervals, and time lags between the physiological assessments and other measures, is an open problem. For example, Ketonen et al. (2022/this issue) aggregated scores for HR and HRV over 5 min to link them to momentary ratings of emotions assessed in experience sampling. This decision has intuitive appeal, especially because the distributions were smoothed using Gaussian weighting. Nevertheless, we cannot be sure if the choice of the interval was optimal. We do not exactly know how respondents judge their current emotional state when responding to ESM. It seems likely that these judgements represent mental aggregations of states over seconds and minutes, but the exact time frame remains unknown and may vary across persons. This problem is exacerbated by the fact that emotion regulation may kick in quickly once any more intense emotion has been triggered—a process that can rapidly change the emotional state, thus producing a mismatch between the intervals captured by physiological and self-report measurement. Research is needed on how to best calibrate the windows of time used to assess physiology, mental processes, and behaviour.

Beyond single assessments, the spacing of measurement occasions also is an open issue. If we want to assess effects of physiological processes on mental processes, or vice versa, then we need knowledge about the temporal unfolding of these effects (for similar considerations in developmental science, see, e.g., Boele et al., 2022). If measurements of cause and effect are too close or too distant, we may miss capturing the effect. If they are too close, then the causal variable may not have sufficient time to change the outcome variable (which can statistically manifest itself in overly high autocorrelations, as noted earlier). If they are too distant, then the effect may already have dissipated.

For example, what is the best time frame to investigate effects of physical activity on students' task-related attention and cognitive performance (Heemskerk et al., 2022/this issue)? Are these effects best captured seconds, minutes, hours, or even days after the activity? Answers may depend on decay rates of both cause and effects (e.g., decay rates of emotions), the immediacy of effects, their accumulation over time, etc. Again, research is needed to find answers to this problem.

THE ROLE OF ENVIRONMENTS AND SOCIO-CULTURAL CONTEXTS

In psychophysiological research, it is typically assumed that physiological processes and their functions are universal across environments and cultural contexts (even if this assumption usually remains implicit). Universality assumptions may be plausible, but they need to be tested empirically. Do physiological arousal and physical activity play the same role in students and teachers from different backgrounds, in different types of schools, and in different socio-cultural contexts? While it may not be possible to answer these questions in any single psychophysiological study given that it is already so expensive to meet the requirements to secure internal validity, it is important to consider the possible role of context in interpreting the findings and to examine their generalizability across contexts in cooperative efforts of multiple labs.

All of the studies reported in this special issue used samples from WEIRD (i.e., Western, Educated, Industrialized, Rich, Democratic) countries, including students and teachers from Finland, Norway, the Netherlands, Italy, Germany, the United Kingdom, and the USA. In future research, we need to move beyond conducting studies in Western countries only and include samples from a broader range of continents and cultural contexts (see also Pekrun & Marsh, 2022).

GUIDELINES FOR FUTURE RESEARCH

To use psychophysiological research to inform educational theory and practice in evidence-based ways, we need to build a cumulative, consistent knowledge base. From the above considerations, a few guidelines can be derived that may help to reach this aim.

1. Status of physiological processes: Specify theoretically and empirically if the target physiological processes should be considered as indicators, components, predictors, causal antecedents, or causal outcomes of mental and behavioural processes.
2. Construct symmetry: Establish construct symmetry of concepts as well as measures of physiological, mental, and behavioural variables.
3. Correlation, prediction, and causation: Once there is sufficient evidence for correlational links, use predictive and causal designs to examine the predictive power and causal status of physiological processes.
4. Within- and between-person designs: Specify the level at which the target mechanisms operate and design the study accordingly. If they operate within persons, then use a within-person design. In case of doubt, also consider the within-person level as physiological mechanisms operate within persons in the first place.
5. Sample size: Include sufficient samples of participants, situations, and measurement occasions. If this is difficult for a single lab, create or join multi-lab studies.
6. Physiological indicators: Include multiple assessments wherever useful to increase reliability, assess multiple physiological indicators to cross-validate the findings for each single indicator, and combine physiological indicators with mental and behavioural indicators when this helps to increase specificity.
7. Time intervals and time lags: Use theory and empirical pilot work to specify the time intervals and time lags that match the duration and decay rates of physiological, mental, and behavioural processes, and the temporal structure of the effects linking them.
8. Social-cultural context: Attend to the role of context in designing and interpreting studies and investigate the generalizability of findings across contexts in broader multi-lab programs of research.

CONCLUSION

Until recently, educational psychology has neglected the importance of physiological processes to understand students' and teachers' motivation, emotions, and engagement. Such a neglect amounts to dis-embodying mental processes. The 12 contributions to this special issue showcase how we can overcome this neglect. Collectively, they demonstrate how closely mental and physiological processes can be intertwined in educational settings. However, the findings also suggest that further progress is needed to reach firm conclusions. To build a coherent science of physiological processes in education, we need to advance conceptual foundations, refine study designs and analytical approaches, and attend to the role of socio-cultural contexts, as proposed in the eight guidelines listed above.

AUTHOR CONTRIBUTIONS

Reinhard Pekrun: Conceptualization; writing – original draft; writing – review and editing.

CONFLICT OF INTEREST

All authors declare no conflict of interest.

ORCID

Reinhard Pekrun https://orcid.org/0000-0003-4489-3827

REFERENCES

Ajzen, I. (1988). *Attitudes, personality, and behavior*. Dorsey Press.

Ajzen, I., & Fishbein, M. (2005). The influence of attitudes on behavior. In D. Albarracín, B. T. Johnson, & M. P. Zanna (Eds.), *The handbook of attitudes* (pp. 173–221). Erlbaum Publishers.

Barroso, C., Ganley, C. M., McGraw, A. L., Geer, E. A., Hart, S. A., & Daucourt, M. C. (2021). A meta-analysis of the relation between math anxiety and math achievement. *Psychological Bulletin*, *147*(2), 134–168. https://doi.org/10.1037/bul000030

Boele, S., Nelemans, S. A., Denissen, J. J. A., Prinzie, P., Bülow, A., & Keijsers, L. (2022). Testing transactional processes between parental support and adolescent depressive symptoms: From a daily to a biennial timescale. *Development and Psychopathology*, 1–15. Advance online publication. https://doi.org/10.1017/S0954579422000360

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review*, *62*(3), 193–217. https://doi.org/10.1037/h0047470

Byers-Heinlein, K., Bergmann, C., Davies, C., Frank, M. C., Hamlin, J. K., Kline, M., Kominsky, J. F., Kosie, J. E., Lew-Williams, C., Liu, L., Mastroberardino, M., Singh, L., Waddell, C. P. G., Zettersten, M., & Soderstrom, M. (2021). Building a collaborative Psychological Science: Lessons learned from ManyBabies 1. *Canadian Psychology/Psychologie Canadienne*, *61*(4), 349–363. https://doi.org/10.1037/cap0000216

de Bruijn, A. G. M., de Greeff, J. W., Temlali, T. Y., Oosterlaan, J., Smith, J., & Hartman, E. (2022). Objectively measured physical activity during primary school physical education predicts intrinsic motivation independently of academic achievement level. *British Journal of Educational Psychology*, e12527. Advance online publication. https://doi.org/10.1111/bjep.12527

Dettweiler, U., Gerchen, M., Mall, C., Simon, P., & Kirsch, P. (2022). Choice matters: Pupils' stress regulation, brain development and brain function in an outdoor education project. *British Journal of Educational Psychology*. Advance online publication. https://doi.org/10.1111/bjep.12528

Fiedler, K. (2011). Voodoo correlations are everywhere – not only in neuroscience. *Perspectives on Psychological Science*, *6*(2), 163–171. https://doi.org/10.1177/1745691611400237

Graham, M. C., Husman, J., Pekrun, R., Villanueva, I., & Christensen, D. (2022). The dynamic experience of taking an examination: Ever changing cortisol and expectancy for success. *British Journal of Educational Psychology*, e12521. Advance online publication. https://doi.org/10.1111/bjep.12521

Hamaker, E. L., Asparouhov, T., Brose, A., Schmiedek, F., & Muthén, B. (2018). At the frontiers of modeling intensive longitudinal data: Dynamic structural equation models for the affective measurements from the COGITO study. *Multivariate Behavioral Research*, *53*(6), 820–841. https://doi.org/10.1080/00273171.2018.1446819

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. P. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, *20*(1), 102–116. https://doi.org/10.1037/a0038889

Heemskerk, C., Strand, S., & Malmberg, L.-E. (2022). Physical activity predicts task-related behaviour, affect and tiredness in the primary school classroom: A within-person experiment. *British Journal of Educational Psychology*, e12523. Advance online publication. https://doi.org/10.1111/bjep.12523

Hoferichter, F., & Raufelder, D. (2022). Biophysiological stress markers relate differently to grit and school engagement among lower- and higher-track secondary school students. *British Journal of Educational Psychology*. Advance online publication. https://doi.org/10.1111/bjep.12514

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, *2*(8), e124. https://doi.org/10.1371/journal.pmed.0020124

Jõgi, A.-L., Pakarinen, E., & Lerkkanen, M.-K. (2022). Teachers' physiological and self-reported stress, teaching practices, and students' learning outcomes in grade 1. *British Journal of Educational Psychology*, e12529. Advance online publication. https://doi.org/10.1111/bjep.12529

Ketonen, E. E., Salonen, V., Lonka, K., & Salmela-Aro, K. (2022). Can you feel the excitement? Physiological correlates of students' self-reported emotions. *British Journal of Educational Psychology*, e12534. Advance online publication. https://doi.org/10.1111/bjep.12534

Kretzschmar, A., Spengler, M., Schubert, A.-L., Steinmayr, R., & Ziegler, M. (2018). The relation of personality and intelligence – what can the Brunswik symmetry principle tell us? *Journal of Intelligence*, *6*(3), 30. https://doi.org/10.3390/jintelligence6030030

Mainhard, T., Donker, M. H., & van Gog, T. (2022). When closeness is effortful: Teachers' physiological activation undermines positive effects of their closeness on students. *British Journal of Educational Psychology*, *92*, 1384–1402. https://doi.org/10.1111/bjep.12506

Martin, A. J., Malmberg, L.-E., Pakarinen, E., Mason, L., & Mainhard, T. (2023). The potential of biophysiology for understanding motivation, engagement, and learning experiences [introduction to the special issue]. *British Journal of Educational Psychology*. Advance online publication. https://doi.org/10.1111/bjep.12584

Mastromatteo, L. Y., Peruzza, M., & Scrimin, S. (2022). Improvement in parasympathetic regulation is associated with engagement in classroom activity in primary school children experiencing poor classroom climate. *British Journal of Educational Psychology*. Advance online publication. https://doi.org/10.1111/bjep.12501

Moors, A. (2017). Integration of two skeptical emotion theories: Dimensional appraisal theory and Russell's psychological construction theory. *Psychological Inquiry*, *28*(1), 1–19. https://doi.org/10.1080/1047840X_2017.1235900

Murayama, K., Goetz, T., Malmberg, L.-E., Pekrun, R., Tanaka, A., & Martin, A. J. (2017). Within-person analysis in educational psychology: Importance and illustrations. In D. W. Putwain & K. Smart (Eds.), *British Journal of Educational Psychology monograph series II: Psychological aspects of education – Current trends: The role of competence beliefs in teaching and learning* (pp. 71–87). Wiley.

Nosek, B. A., & Lakens, D. (2014). Registered reports: A method to increase the credibility of published results [editorial]. *Social Psychology*, *45*(3), 137–141. https://doi.org/10.1027/1864-9335/a000192

Orth, U., Clark, D. A., Donnellan, M. B., & Robins, R. W. (2021). Testing prospective effects in longitudinal research: Comparing seven competing cross-lagged models. *Journal of Personality and Social Psychology*, *120*(4), 1013–1034.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, *18*(4), 315–341. https://doi.org/10.1007/s10648-006-9029-9

Pekrun, R. (2009). Global and local perspectives on human affect: Implications of the control-value theory of achievement emotions. In M. Wosnitza, S. A. Karabenick, A. Efklides, & P. Nenniger (Eds.), *Contemporary motivation research: From global to local perspectives* (pp. 97–115). Hogrefe.

Pekrun, R. (2020). Self-report is indispensable to assess students' learning. *Frontline Learning Research*, *8*(3), 185–193. https://doi.org/10.14786/flr.v8i3.637

Pekrun, R., & Marsh, H. W. (2022). Research on situated motivation and emotion: Progress and open problems. *Learning and Instruction*, *81*, 101664. https://doi.org/10.1016/j.learninstruc.2022.101664

Pekrun, R., Marsh, H. W., Elliot, A. J., Stockinger, K., Perry, R. P., Vogl, E., Goetz, T., van Tilburg, W., Lüdtke, O., & Vispoel, W. (2023). A three-dimensional taxonomy of achievement emotions. *Journal of Personality and Social Psychology*, *124*(1), 145–178. https://doi.org/10.1037/pspp0000448

Roos, A.-L., Goetz, T., Krannich, M., Donker, M., Bieleke, M., Caltabiano, A., & Mainhard, T. (2022). Control, anxiety and test performance: Self-reported and physiological indicators of anxiety as mediators. *British Journal of Educational Psychology*. Advance online publication. https://doi.org/10.1111/bjep.12536

Scherer, K. R., & Moors, A. (2019). The emotion process: Event appraisal and component differentiation. *Annual Review of Psychology*, *7*, 719–745. https://doi.org/10.1146/annurev-psych-122216-011854

Schönbrodt, F., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, *47*(5), 609–612. https://doi.org/10.1016/j.jrp.2013.05.009

Schultzberg, M., & Muthén, B. (2018). Number of subjects and time points needed for multilevel time-series analysis: A simulation study of dynamic structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(4), 495–515. https://doi.org/10.1080/10705511.2017.1392862

Törmänen, T., Järvenoja, H., Saqr, M., Malmberg, J., & Järvelä, S. (2022). Affective states and regulation of learning during socio-emotional interactions in secondary school collaborative groups. *British Journal of Educational Psychology*. Advance online publication. https://doi.org/10.1111/bjep.12525

van Vemde, L., Donker, M. H., & Mainhard, T. (2022). Teachers, loosen up! How teachers can trigger interpersonally cooperative behavior in students at risk of academic failure. *Learning and Instruction*, *82*, 101687. https://doi.org/10.1016/j.learninstruc.2022.101687

Vansteenkiste, M., Ryan, R. M., & Soenens, B. (2020). Basic psychological need theory: Advancements, critical themes, and future directions. *Motivation and Emotion*, *44*, 1–31. https://doi.org/10.1007/s11031-019-09818-1

Voelkle, M. C., Brose, A., Schmiedek, F., & Lindenberger, U. (2014). Toward a unified framework for the study of between-person and within-person structures: Building a bridge between two research paradigms. *Multivariate Behavioral Research*, *49*(3), 193–213. https://doi.org/10.1080/00273171.2014.889593

Weiner, O. M., & McGrath, J. J. (2017). Test-retest reliability of pediatric heart rate variability. A meta-analysis. *Journal of Psychophysiology*, *31*(1), 6–28. https://doi.org/10.1027/0269-8803/a000161

Zaccoletti, S., Raccanello, D., Burro, R., & Mason, L. (2022). Reading with induced worry: The role of physiological self-regulation and working memory updating in text comprehension. *British Journal of Educational Psychology*. Advance online publication. https://doi.org/10.1111/bjep.12491

How to cite this article: Pekrun, R. (2023). Mind and body in students' and teachers' engagement: New evidence, challenges, and guidelines for future research. *British Journal of Educational Psychology*, *93*(Suppl. 1), 227–238. https://doi.org/10.1111/bjep.12575