



The journey to proficiency: Exploring new objective methodologies to capture the process of learning and professional development

Christian Harteis, Ellen Kok, Halszka Jarodzka

Introduction

Over the last decades, educational research established different foci on learning – all of them aiming at understanding how learning takes place and how its outcomes can be improved by instruction. They can be distinguished regarding the object of observation: (a) There is research focusing input to learning processes, particular discussing content and teacher or learner characteristics influencing learning; (b) other research approaches exist that focus learning processes themselves, (c) finally there is research focusing learning outcomes. What varied over the time is the emphasis of these foci. For example, the current interest on international comparisons of educational systems (i.e. TIMSS, ISGLU, PISA) addresses research with focus on learning outcomes.

Within the last two decades, the focus of research on learning and instruction shifted from an emphasis on the outcomes back to the processes that underlie learning. However, in contrast to process research between 1960ies and 1980ies that investigated school-based classroom teaching (Shulman, 1986), current process research spreads across the entire lifetime and across all skill levels, from kids studying a textbook for 20 minutes to professionals developing extraordinary expertise over years. How varying these fields may seem, their mutual aim is to understand the mental structures and their changes through learning including social, motivational, and emotional aspects influencing learning processes. However, for decades it was extremely challenging to capture these processes in a meaningful manner.

Recent development of software methodology and hardware technology opened fascinating opportunities for educational research. On the one hand, the development of data analysis methodology, such as machine learning or data sequence comparison and string detection, is reaching a level so that it can be applied to approach new research questions within educational science. On the other hand, a variety of sensors made various measurements originating from highly specialized fundamental research (e.g., electroencephalography, cardiovascular measures, infrared eye-tracking) seemingly accessible and applicable. For instance, in the early 2000's companies started building 'plug & play' eye trackers with ready-to-use analysis software that claimed to guide its users intuitively. More recently, relatively cheap wearable devices for measuring brain activity (EEG) and electrodermal activity have become available. It is the combination of decreasing prices for and sizes of sensor systems with increasing usability of operating and analysis software and the development of novel, easier to apply analysis methodologies that reduced inhibition threshold for application the area of education.

Hence, many researchers started utilizing these methodologies in a wide area of educational research. However, quickly it turned out that neither was the use of the hardware as easy as the sellers claimed, nor was the analysis of the data as straightforward. All online measures of learning create a kind of data that is not comparable with traditional qualitative or quantitative empirical data. Many online measures collect longitudinal data in a frequency of milliseconds and, thus, generate thousands of data-values. While researchers found many opportunities these measures offer, they also faced many challenges. These comprise a variety of problems, e.g. detecting meaningful events in high-frequency measures, combining process measures of different granularities, synchronizing measures, capturing the sequential nature of learning processes and defining reasonable frequencies for statistically analyzing skewed, multilevel data sets. What keeps happening, though, is that researchers face the same issues or similar problems on and on as they cannot get easily access to the progress made by others in their field. Due to this lack of exchange, researchers often have to re-invent the wheel. On top of that, online measures of learning operate on a granularity that does not easily match with the predictions that can be made from our current theoretical models of learning and expertise development. Online measures provide data on a very micro level of learning whereas theoretical models usually address the macro level of development. Thus, researchers that use process measures are in need of ways to exchange thoughts, not just about methodological issues, but also how methodological choices relate to theoretical models.

Hence, researchers started activities to establish more or less formal groups which aim at sharing their experiences (e.g., one of these is the EARLI SIG 27 'online processes of learning'). However, also within already existing communities the interest for these measures rose, such as in the EARLI SIG 14 'learning and professional development'. We argue that this exchange is crucial for meaningful and fruitful further development within educational research, in particular as the usage of such techniques is growing.

Just as an example, the methodology of eye tracking is extremely growing with over 700.000 Google scholar hits in the past decade (~370.000 in 1998 – 2007, ~ 66.000 in 1988 – 1997, and less than 30.000 publications before 1987).

Therefore, it is important to go beyond purely talking about experiences with process measures. What is needed is explorative, methodology-focused research in order to initiate negotiation within the scientific community of educational researchers on how these novel approaches of data-collection and data-analysis contribute to the communities' state of knowledge on learning and instruction. Thus far, the challenges of using process measures go by unnoticed as it is hardly possible to discuss them in traditional empirical study papers, which focus on knowledge structures and their changes through learning instead of on methodological developments. The current special issue in *Frontline Learning Research* is a first step to fill this gap.

About this special issue

This special issue aims at exploring possibilities of using process data on learning in different contexts and critically discussing exactly these methods with respect to their explanatory power for learning and expertise development, and for gaining insight on its underlying processes operating within cognitive structures. These contributions present experiences in applying new methodologies and put the findings up for discussion. The goal of all contributions is to reflect the strengths and limitations of their measures and to provide a statement on how informative their data can be for researching learning. It addresses the broad readership within the EARLI community. The idea for this special issue resulted from two well attended and highly appreciated SIG-invited symposia (SIG 14 and SIG 27) for the EARLI conference at Tampere, Finland in 2017. A public call for contributions provided additional contributions and we invited three discussants to reflect upon the articles. Now, more than one year after this conference, this special issue provides a broad set of papers reporting and reflecting selected approaches of online measures of learning processes. We intentionally considered contributions from various fields and domains of educational research. The EARLI community represents educational research covering the entire life-span and applying laboratory conditions as well as conducting field research. All contributors were encouraged to discuss carefully if and how the selected methodological approaches and data can be informative not only for the context of the respective paper but for the educational research community in general. Hence, we hope that everybody can find novel, interesting and fruitful information within this special issue.

The methods discussed in this Special Issue cover neuroscience topics, such as the possibilities of neuroscience for education (Van Atteveldt et al., this issue), EEG (Scharinger, this issue). It also discusses different applications of eye tracking, such as machine learning to analyze eye tracking data (Garcia Moreno-Esteva et al., this issue; Harteis et al., this issue), its limitation in investigating web search processes (Salmeron et al., this issue), or the challenge of combining it with musical performance (Puurtinen et al., this issue). Further contributions discuss the use of physiological data, such as skin conductance (Eteläpelto et al.; Nokelainen et al.; both this issue) and combine it with further measures to study collaboration (Hoogeboom et al., this issue), logfile analyses to study blended learning (Van Laer & Ellen, this issue), or even prosodic analyses of conversations in classrooms (Hämäläinen et al., this issue). Several contributions use specifically multimodal aspects to study collaboration (Hoogeboom et al., this issue), observational data for analyzing teachers' behavior (Donker et al., this issue), or self-regulated learning (Järvenoja et al., this issue).

Overarching themes

Taken together, the contributions to this special issue discuss opportunities and limitations of process measures, their combination with each other, or their combination with conventional measure for analyzing learning processes. They reveal the following overarching challenges.

Objective data

One important benefit of process measures as presented here, is that they do not rely on self-reports, and, thus, can be considered as objective data. We must keep in mind, though, that their interpretation remains subjective to the experimenter. Furthermore, these methods open opportunities to gather data about unconscious regulation processes, whereas self-reports necessarily provide access only about what participants are aware of. It is important to keep in mind, however, that sometimes subjective data is more appropriate for a given research question. A clear understanding of what type of data would fit a certain research question remains the crucial challenge of utilizing online measures.

Multimodal data

As several contributions to this special issue reveal, one important development is the increased use of multimodal designs. Such designs aim to gather a broader understanding of real-world learning situations by using different types of data (e.g., a combination of psychophysiological data, video data and data from sociometric badges). Severe challenges hereby are that the added value from each of those types of data for the research question needs to be clear, that the data is often collected at different levels (e.g., some data was collected at the team level and other data was collected at the individual level), as well as at different sampling frequencies (e.g., eye-tracking data can be measured at a level of 500 Hz, skin conductance levels are measured at 4 Hz, whereas team effectiveness is only measured once). Synchronization of data can be challenging, both on a practical level (i.e., data files should use the same time representation) as well as on a sampling level (i.e., data from different participants need to be synchronized).

An important aspect of many of these process measures is that they are potentially personal data (i.e., they may be traced back to one specific person). This is particularly the case if several data sets are collected from one participant. In such a case, the new

GDPR regulations come into play (<https://eugdpr.org>). We should also keep in mind that even if a process data set is not yet easily traceable back to one person, it may easily become so with the fast development of machine learning techniques in near future.

Analysis

Those large data-sets also call for an improvement in our techniques of analysis, as the current conventional statistics do not allow for a non-biased analysis of large numbers of variables at the same time. Furthermore, data is often averaged or summed up over time, so the temporal order of data might get lost, while the process of interest could be reflected in the temporal order. Data mining and machine learning techniques are one important option that apply complex algorithms based on different mathematical models than inferential statistics. They provide novel opportunities of revealing hidden patterns within huge data sets. Hence, they allow for testing the predictive value of sets of variables for certain outcome measures and, thus, make it possible to quantify and statistically test which online measures predict the outcome.

At the same time, a careful theory-driven decision as to what variables should reflect the processes of interest is still critical, and the fact that large numbers of variables are available should not tempt us to simply report large numbers of analyses, and cherry-pick the interesting (significant) effects for discussion.

Ecological validity

Many of the authors argue that their designs allow for collecting data in the field (instead of laboratory environments), and, as such, increases ecological validity. The underlying assumption is that ecological validity, the extent to which the study approximates 'the real world', predicts external validity, the extent to which the study generalizes to 'the real world'. In particular because many of the measures do not disturb natural task performance. However, studies within this special issue showcase that ecological validity does not necessarily translate into external validity (e.g., fluctuations in skin conductance under 'natural' conditions might also reflect body posture or movement instead of mental effort). It depends on the context and the research interest if and how fuzzy data appear acceptable or precise data are required. We can find argumentations that a compromised quality of, e.g., easy-to-use EEG apparatus is acceptable in the context of the new opportunities for neuroimaging in 'real-world' settings, and we can find argumentations that low data quality and the confounding effects that could occur in ecological valid environments may occlude real effects and, thus, compromise external validity.

Hence, it is often useful to start from hypotheses generated by laboratory research and investigate these in real-world settings. However, the opposite could also provide insights: taking important but tentative findings from field studies and bringing these into the lab (including, where possible, the real-life complexity) to understand the mechanisms involved in more detail.

More ecologically valid testing environments are mostly useful when appropriate analysis techniques are available. The fixation-related EEG frequency band power analysis that Scharinger introduced (in this special issue), for example, is an analysis technique that makes it possible to investigate multimedia learning environments using EEG, where previous EEG research on reading required the presentation of single words instead of free reading tasks.

Coupling of high-level theoretical models to fine-grained data streams

Common theoretical concepts of learning and development describe processes that usually last (much) longer than milliseconds. Online measures, however, bear the particular quality to provide data on a very high resolution. It is important to keep the different granularity in mind when developing research questions and designing research settings. There may be theoretical frameworks that require less precision than others: Understanding visual expertise and pattern recognition, on the one hand, focuses detailed phenomena of physical behavior that may remain under the surface of consciousness. This case requires quite a precise coupling of theories of vision with the data-streams. On the other hand, investigating the importance of emotions for learning may allow more fuzziness, as long as the duration of an emotional state is considered less important for learning processes than the pure occurrence of an emotion. Hence, there is neither 'the' challenge of coupling theories with data, nor is there the 'one fits for all' solution. In the developing field of researching learning with process data the full breadth of opportunities can be found. What concretely is to be considered crucially depends on necessary theoretical decisions.

Conclusions and outlook

All contributions have put their findings up for discussion and reflected on strengths and limitations of the measures applied in their studies. As such, this special issue provides a valuable resource for any researcher who already works with process measures or start working with process measures. Based on the contributions, we derived suggestions on how to implement new methods and technologies to our applied field of educational science in a meaningful way.

Choosing a method

When thinking about using a new method or technology for research, we always must ask ourselves, why we need it. Is it to address an otherwise not possible to address hypothesis? Is it to explore thus far hidden processes? Or is it rather to simply try out a fancy, new technology that was thrown at us? Whichever the answer may be, we need to be clear about it. Based on the experiences gathered in this special issue, we strongly advice to always consider how this new methodology or technology will help to approach

the research question. Moreover, we recommend to be cautious when being drawn towards new gadgets out of pure curiosity and we advise to stay always as low tech as possible and as high tech as necessary.

Implementing a methodology When researchers start to use a new method, it is critical that they understand where the method comes from, what its history is (i.e., in which fields is it already successfully applied and how?). This often helps to understand why certain approaches were chosen and decisions were made. How is this technique currently used in its 'home' domain? How are the experiments set up? What are the analysis techniques? Even though many of the 'new' methods are (relatively) new to educational research, often there is a large body of research available in other domains, that can inform the researchers. It is important to get to know this field to make sure that no huge mistakes are made because you do not know the field, as this might result in invalid data recording or analysis. We suggest to always begin by cooperating with an expert from the original field.

Analysis

Since computers and software developed similar rapidly further, it is now possible to utilize calculator power and software algorithms for completely new procedures of data analysis (e.g., big data, data-mining). Since we, as educational researchers, might not have the expected background in statistics, computer science or data-science to execute those kind of procedures, it is central that we collaborate with researchers and practitioners from other fields. The combination of different types of expertise is central for progress in the use of process measures.

Interpretation

An important challenge is how to incorporate these new methodologies that measure fine grained processes with our theories that make statements on more macro levels. How to derive predictions from our theories to these measures and how to make meaningful statements for our theories from these empirical findings? These questions should not only guide our choice of research methodology, but also challenge us to further develop and specify existing theories and frameworks.

Drawing conclusions

To conclude, we hope that this special issue provides a starting point for more methodological papers in educational sciences, which critically discuss the application of (new) technological approaches and process measures, including validity information for that (set of) measure as well as practical advice for their use.

This field is developing rapidly, so we also tried to realize this special issue quite swift in order to contribute to the start of the scientific discourse on those issues. We are aware though, that the development just started and we are far away from fully understanding potential and limitations of these new kind of data and measurements. Hence, we hope to see more methodological publications discussing new ways of capturing learning processes in the future!

Acknowledgements

The guest editors would like to thank the reviewers of the special issue.

References

Shulman, L. S. (1986). Paradigms and research programs in the study of teaching. In M. Wittrock (Eds.), *Handbook of research on teaching* (pp. 3-36). New York: MacMillan.