Frontline Learning Research
An official journal of EARLI

Frontline Learning Research, Vol 6 No. 3 Special issue (2018) 6-36
ISSN 2295-3159

A multi-componential methodology for exploring emotions in learning: using self-reports, behaviour registration, and physiological indicators as complementary data

Anneli Eteläpelto[a], Virpi-Liisa Kykyri[a], Markku Penttonen[a], Päivi Hökkä[a], Susanna Paloniemi[a], Katja Vähäsantanen[a], Tuomas Eteläpelto[b], &Vesa Lappalainen[b]

a Faculty of Education and Psychology, University of Jyväskylä, Finland

b Faculty of Information Technology, University of Jyväskylä, Finland

Article received 22 May 2018 /revised 30 August /accepted 1 October /available online 7 December

Abstract

Studies on emotions in learning are often based on interviews conducted after the learning. These do not capture the multi-componential nature of emotions, nor how emotions are related to the processes of learning. We see emotions as dimensional, multi-componential responses to personally meaningful events and situations. In this methodologically advanced pilot study we developed a multi-componential methodology, capable of providing complementary information on emotions in professional learning. For this purpose, we used a within-subject design applied to a single individual, with a focus on emotions during professional learning. Within a laboratory setting, the subject was shown personally meaningful video extracts from a learning situation in which she had previously participated. The data were gathered through (i) self-reports of emotions via the Emotion Circle (EC) online assessment tool, (ii) measures of autonomic nervous system (ANS) activity obtained via Electrodermal Activity (EDA) and Heart Rate Variability (HRV), (iii) behavioural registration of facial expression and gaze, and (iv) the Stimulated Recall Interview (SRI). Self-reports of emotions via EC, and also the emotion-driven SRI, were found to be productive, not only in detailing and explaining emotions experienced during the viewing of the videos, but also in bringing about reflective learning and novel insights. EDA and HRV provided complementary information on the subject's ANS activity during the learning process. We present conclusions and future challenges in applying a multi-componential methodology to research emotions within professional learning.

Keywords: learning process; emotions; multimethod measuring; on-line self-reports; autonomic nervous system

Corresponding author: Anneli Eteläpelto, P.O. Box 35, 40014 University of Jyväskylä, Finland anneli.etelapelto@jyu.fi
DOI: https://doi.org/10.14786/flr.v6i3.379

Acknowledgment

Professor Anneli Eteläpelto's highly-meritorious professional career in the field of adult education at University of Jyväskylä (Finland) will continue in her new role as an emerita. We greatly appreciate her years of dedication to research, and especially all the inspiring ideas she has introduced over the years! Among other roles, she has served the academic community as the coordinator of EARLI SIG14 "Learning and Professional Development" and hosted the SIG meeting in 2008. Further, she is renowned for her repeated successes in securing prestigious grants from the Academy of Finland and elsewhere. Anyone who has had been fortunate enough to collaborate with Anneli has probably found her an inspiring colleague and, sometimes, surprising in her contributions. In her new role as a professor emerita, she will undoubtedly continue to contribute actively to the discourse on learning and professional development. This special issue describes a starting point for new direction of research, which she has actively advocated, and we hope she will remain involved in these developments also in the future to see the vision for which she advocated fully realised.

All the best and warm regards on behalf of all your colleagues!

Christian, Raija, Erno, and Stephen

1. Introduction

There is fairly convincing evidence on the vital role of emotions in learning. Emotions are related to motivational processes, self-efficacy, and active engagement, each of which has a salient role in productive learning in schooling contexts (Pekrun, Elliot & Maier, 2006). In research on adult learning, positive emotions have been found to broaden the scope of perception, whereas anxiety and fear have been connected to a narrowing of the perception and curiosity necessary for active and agentic learning (Fredrickson & Branigan, 2002; Hökkä, Vähäsantanen, Paloniemi & Eteläpelto, 2017; Perry, 2006; Storbeck & Maswood, 2015; Sung & Yih, 2015). In team-based learning, social and self-conscious emotions such as compassion, love, shame, anxiety, and anger have been found to influence how team members see each other, and how they perceive the future of the team (Homan, Van Kleef & Sanchez-Burks, 2015). Furthermore, it has been shown that in work organizations emotions critically influence the work-related learning manifested in job performance, motivation, creativity, decision-making, turnover, psychological wellbeing, teamwork, and leadership (Barsade & Gibson, 2007).

Despite convincing findings on the vital role of emotions in human learning, there has been a lack of research on the role of emotions in professional learning settings. These are characterized by processes of active influencing and developing, and by the negotiation of professional identity (Eteläpelto, Vähäsantanen, Hökkä & Paloniemi, 2013; Vähäsantanen, Räikkönen, Paloniemi, Hökkä & Eteläpelto, 2018). Recent studies have emphasized that professional learning occurs, in particular, via collaboration within social interaction, plus experimentation and reflection regarding one's professional mission and practices (Zwart et al., 2015). Moreover, to achieve useful outcomes, social learning processes should be designed in such a manner that professionals have sufficient resources to shape their own professional identity and work (Eteläpelto, Vähäsantanen, Hökkä & Paloniemi, 2014). This touches on the notion that if meaningful learning is to be achieved, it is crucial to acknowledge and promote agency, and negotiations on professional identity (Philpott & Oates, 2017).

Here we emphasize reflection and insight concerning one's own ways of thinking and interacting as important means of professional learning, after particular events. This kind of learning could manifest itself, for example, in insights into one's ways of thinking and acting, plus the reasons for these, with greater overall self-awareness emerging as a result (Vall et al., 2018). Note that we have here adopted the term 'professional learning' as denoting the individual as an *active and reflective participant*, that is, someone who is responsible for learning and for constructing change at a personal level within a given context (Labone & Long, 2016; Eteläpelto, 2017; Vähäsantanen, Paloniemi, Hökkä & Eteläpelto, 2017a).

Recent studies have shown that processes of group-based identity learning are particularly imbued with strong emotions, and that these emotions strongly influence actual learning outcomes, manifested as renegotiated work identities (Vähäsantanen, Hökkä, Paloniemi & Eteläpelto, 2017b; Vähäsantanen, Hökkä, Paloniemi, Herranen & Eteläpelto, 2017). However, the findings on emotions are based on interview data collected *after* the learning experiences. Thus, they do not take into account the multi-componential nature of the emotions (involving subjective experience, the autonomic nervous system, and behavioural changes), or how these are related to the processes of learning. A truly multi-componential understanding of emotions would appear to require a *multi-componential approach to the measuring of emotions*. The question then arises of how to overcome the methodological challenges of multi-componential measurement, and how to apply methodologically advanced tools to investigate emotions and related learning. In fact, there are now unobtrusive, technologically advanced tools such as face reading and gaze analysis (e.g. Azevedo et al., 2013, 2016; Zembylas & Schultz, 2016) which, together with psychophysiological and self-assessment methods, can be used in the simultaneous collection of multiple data from different subsystems of emotions. However, in such simultaneous measurements, one must address questions of complementarity, interchangeability, validity, and reliability. In addition, there are challenges regarding the differing time windows of the various modalities.

In addition to evidence on the relationships between learning and emotion, there has long been discussion on how emotions trigger or inhibit human behaviour (Hommel, Moors, Sander & Deonna, 2017), and how the emergence of emotions is closely connected to autonomic nervous system (ANS) activity (Levenson, 2014; Mauss et al., 2005). Recent discussion on emotion has been critical of mind-body dualism, pointing to evidence on how closely these interact, especially in the affective domain. It has long been recognized that ANS has a central role in emotions. As Levenson (2014, p. 100) has noted, 'when it comes to emotions, all roads lead to ANS'. The ANS functions via two opposite but interacting regulation systems, i.e. the *sympathetic* and the *parasympathetic* nervous system. In experiences of fear and anxiety, the sympathetic nervous system produces the 'fight or flight' response, whereas the parasympathetic nervous system is responsible for calming our body and mind ('rest and digest'). The ANS is thus closely allied to the human experience of emotions, and the two branches of the ANS are closely intertwined with behavioural responses, including changes in facial expressions and the active focusing of attention. It should be noted that the multi-componential nature of emotions has been accepted in a wide range of theoretical models of emotions (Kreibig, 2010; Mauss & Robinson, 2009; Gendolla, 2017).

Since early research on emotions, there has been a consensual understanding on emotions as comprising the following subsystems: (i) subjective experiences, (ii) psychophysiological responses emerging from the unconscious functioning of the ANS, and (iii) behavioural and action-related manifestations. Although there is agreement on the multi-componential nature of emotions, there is no agreement on the relations between the components of subjective experience, the ANS, and behavioural changes. Cognitive theories of emotions assume that there is a top-down relationship between levels in the subject's cognitive assessment of emotions, and that this influences the responses of the ANS and behavioural phenomena (e.g. Lazarus, 1991). In contrast with this, evolutionarist-functionalist theories view emotions as organizing the activity of the ANS and other physiological systems (Levenson, 2014). The relationships between the different subsystems are often addressed in terms of coherence, referring to the coordination, or association, of a person's experiential, behavioural, and physiological responses as an emotion unfolds over time (Mauss et al., 2005). Despite active theoretical discussion of such coherence, and of the nature of relationships between the three response systems, there is a lack of empirical evidence concerning how emotions organize activity within the ANS, and between the ANS and other response systems, including facial expression and subjective experience (Levenson, 2014). Nevertheless, empirical research has revealed that subsystem coherence and synchronization vary between different emotions (Kreibig, 2010; Levenson, 2014). It has also been found that different emotions (such as fear and joy) may activate different patterns of ANS activity. However, research does not agree on the patterning, coherence, and specificity of the different subsystems (Levenson, 2014).

As indicated above, agreement on the multi-componential nature of emotions implies multi-method measurement of emotions. However, disagreement on the direction of influence between mind, behaviour, and ANS responses brings challenges in interpreting the relationships between data derived from different subsystems. This implies a need for a pilot study which would address the relevant methodological challenges. We therefore sought to construct a multimethod research setting within which we could collect information

simultaneously from the different subsystems of emotions operating within professional learning. In this article we consider the kinds of information that appear to emerge from such a methodology. Our aim is to form some tentative conclusions on how the methods used may provide *complementary data*, allowing us to capture different aspects of emotions, plus their role within work-related learning processes.

The study reported here was based on an understanding of emotions as multi-componential phenomena, manifested simultaneously in subjective experience and in the unconscious autonomic nervous system, and also in behaviour (with nonverbal behaviour operating somewhere in between these, being partly conscious and partly unconscious). The study aimed to capture emotions simultaneously on three levels, namely (i) subjective experience, (ii) ANS, and (iii) behaviour. The methods can be summarized as follows (see also Figure 2, Section 5.1):

(i) Video-recorded episodes of previous learning situations were presented to the subject here referred to as 'Lisa' (pseudonym) in a laboratory setting. Her *subjective experiences* of these were elicited using an on-line application called *Emotion Circle* (EC), developed for the self-assessment of emotions concurrently with watching videos (see Section 5.2.1). The videos were of episodes from a training programme in which Lisa had participated previously. The episodes were selected on the basis that Lisa had perceived them as personally meaningful, and as containing highly significant events for her learning within the program. The concurrent assessments of emotions (plus their connections to learning) were elaborated using the *Stimulated Recall Interview* (SRI) method (Kagan, Krathwohl & Miller, 1963; Kykyri et al., 2017). The SRI interview took place immediately after Lisa had given her EC assessment.

(ii) ANS *activity* was measured using one of the most reliable indicators of arousal emerging from sympathetic nervous system activity, namely *Electrodermal Activity* (EDA), as measured via *Skin Conductance* (SC). In addition, *Heart Rate* (HR) and *breathing* were registered as indicators of ANS activity involving the parasympathetic nervous system, manifested as *Heart Rate Variability* (HRV).

(iii) *Behavioural indicators*, which are partly regulated by the autonomous nervous system, include facial expressions. These were captured through non-intrusive video-recordings, capable of being analysed via the computational methods of FaceReader (FaceReader Version 6.1., 2015). In addition, the focus and direction of visual attention were registered using gaze recording (BeGaze Version 3.7. Manual, 2016).

In this single case study, a within-subject design was used. A within-subject design is recommended by researchers investigating concurrent responses and the coherence of multiple sub-systems of emotions (Levinson, 2014; Mauss & Robinson, 2009; Kreibig, 2010). Our purpose was to pilot the research procedure and multi-componential methodology required when one is investigating concurrent measures applicable to changing emotions, especially those emotions that occur within professional learning processes. In the best case, this would give us insights into the subjective experience of emotion – observed concurrently with ANS functioning, and the behavioural manifestations of emotions – plus indications on how the measures may provide information on emotions in learning. We anticipated that the study conducted on 'Lisa' would contribute to an understanding of the kinds of complementarity that can occur between the indicators used and the three different systems under study (subjective experience, the ANS, and behaviour).

The following section specifies what is known about emotions in terms of (i) the sub-components of emotion, (ii) dimensional vs. discrete models, and (iii) how emotions are related to action. Based on this specification, we present our understanding and definition of emotions, necessary for addressing the validity of the methods used.

2. The conceptualization of emotion

From Darwin (1872/1965) onward, researchers on emotions (Ekman, 1992; Lazarus, 1991; Levenson, 1994) have argued that emotions involve coordinated changes across experiential, behavioural, and physiological response systems (Mauss et al., 2005). Tomkins (1962) has suggested that emotions are sets of

organized responses that (when activated simultaneously) are capable of simultaneously capturing widely distributed organs (the face, the heart, and the endocrines) and imposing on them a specific pattern of correlated responses. Most, but not all, of the theorists above have taken a functional perspective, proposing that by imposing coherence across response systems, emotions facilitate the organism's response to environmental demands, and prepare the organism for a set of diverse actions (Levenson, 1994, 2003; Mauss et al., 2005).

For many theorists, a defining feature of emotion is response system coherence (Mauss et al., 2005). This refers to the coordination, association, concordance, or organization of the response tendencies pertaining to a person's experiential, behavioural, and physiological responses as the emotion unfolds over time (Ekman, 1992; Lazarus, 1991; Levenson, 1994; Scherer, 2009; Tomkins, 1962). The notion that response coherence is a core feature of emotion suggests two corollaries. First of all, response coherence should increase as the intensity of emotion increases. Weak emotions may provoke little coordination of the response systems, whereas strong emotions may provoke greater coordination. Secondly, different emotions should be associated with different patterns of experiential, behavioural, and physiological response, tailored to meet the demands of different situations (Mauss et al., 2005). For instance, *amusement* might be associated with facial displays of amusement, increased body activity, and a commensurate pattern of increased cardiovascular and electrodermal responding. By contrast, *sadness* might be associated with facial displays of sadness, decreased body activity, and a commensurate pattern of decreased cardiovascular and electrodermal response (Mauss & Robinson, 2009).

Nevertheless, in a review by Mauss and Robinson (2009), the researchers observed that, in contrast with the theoretically assumed coherence of the response systems, the empirical findings have been mixed. In fact, psychophysiologists have long emphasized the weak correlations between experiential and physiological response systems, and even between various measures within the physiological response system. Along similar lines, more recent studies have found relatively modest correlations between experiential, behavioural, and physiological measures in the context of specific emotional states such as fear (Mauss et al., 2005).

In recent discussions on emotions, there has been a debate between discrete and dimensional models of emotions. Discrete models have assumed that different emotions are associated with discrete and invariant patterns of response within each response system. By contrast, dimensional models assume that measures of emotional response reflect dimensions rather than discrete states. Dimensional perspectives argue that emotional states are organized by underlying factors such as valence and arousal (e.g. Barrett & Russell, 1999). Discrete emotion perspectives, by contrast, suggest that each emotion (e.g. anger, sadness, joy) has unique experiential, physiological, and behavioural correlates. Having conducted a review, Mauss and Robinson (2009) concluded that research tends to support the dimensional perspective. Thus, the recent consensual, componential model of emotions conceptualizes emotions as dimensional phenomena comprising experiential, physiological, and behavioural responses to personally meaningful stimuli (Mauss & Robinson, 2009; Russell, 2005). In such a model (see Figure 1), an emotional response begins with appraisal of the personal significance of an event (Lazarus, 1991), which in turn gives rise to an emotional response involving subjective experience, physiology, and behaviour (Frijda, 1986).

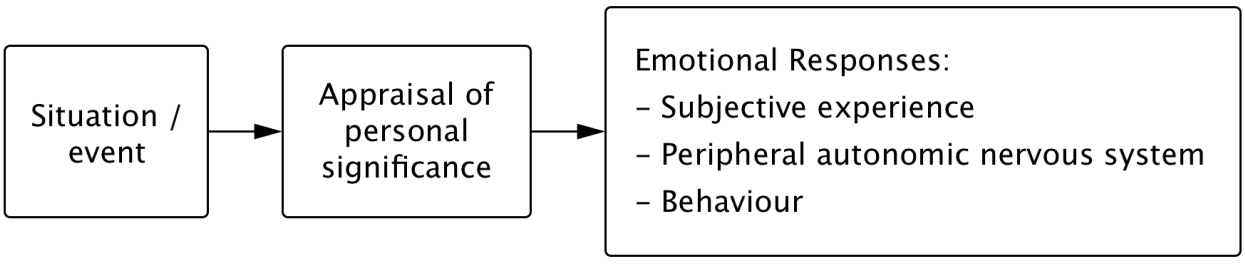

Figure 1. A consensual component model of emotional response (modified from Mauss & Robinson, 2009)

A dimensional model of emotions is also present in the *circumplex* model of emotion (Russell, 2005), which we have used as a framework in developing the *Emotion Circle* (see Section 5.2.1). This is a tool aimed at the on-line assessment and self-reporting of emotions (Eteläpelto et al., 2017). The circumplex model proposes that all affective states arise from cognitive interpretations of core neural sensations, which are the product of two independent neurophysiological systems. In the circumplex model, the vertical dimension depicts the level of arousal, whereas the horizontal dimension depicts the *valence* (pleasure vs. displeasure) of emotions. Posner, Russell and Peterson (2005) argue that the circumplex model of affect is consistent with many recent findings from behavioural, cognitive neuroscience, neuroimaging, and developmental studies of affect.

Over many years there has been discussion on whether and how emotions influence action, including active learning. In a summary of recent research, Hommel, Moors, Sander, and Deonna (2017) conclude that there is little doubt that emotions influence action. Emotions can influence the motivation process, and thus action, by fulfilling at least three functions. First of all, the emotions experienced can function as strong need-like motivational states. Secondly, anticipated emotions can function as incentives, and justify action. Thirdly, emotions can give information on progress in goal pursuit, permitting behaviour calibration (Hommel et al., 2017).

Although the various theories assume a causal relation between emotion and action, they nevertheless disagree on the direction of this causal relation. Feeling theories, such that of James (1884), see action as partly the cause of emotion. By contrast, other theories take emotion to be the cause of action. Unsurprisingly, theories that see actions as part of emotions expect the processes involved in generating emotions to play a role in action as well (Hommel et al., 2017).

In the present article, emotions are understood as dimensional and multi-componential (i.e. experiential, physiological, and behavioural) responses to a personally meaningful antecedent event or situation, causing changes in the quality of subjective feeling, expressive behaviour, and physiological activation (Kreibig, 2010; Levenson, 2014; Mauss & Robinson, 2009). The multi-componential and dimensional understanding of emotions does not imply coherence between the different subsystems of emotion. Neither does it specify the causal relationships that may exist between changes in the three components of emotions. However, coherence might increase as the intensity of the emotion increases. A dimensional understanding means that major variation in the subsystems of emotions may occur in terms of intensity and valence, without separate emotions necessarily having a specific fingerprint within or across different subsystems. Furthermore, we assume that emotions are closely intertwined with the active processes of learning, even if the direction of influence or the specific causal links cannot be hypothesized.

A lack of coherence between the subsystems of emotion can be manifested for several reasons. For example, subjective experiences of emotions can change without changes in autonomic nervous system responses, and ANS changes may occur without concomitant changes in subjective experience. The reasons for this could derive, for example, from the lack of a subject's faithful report on the emotional states experienced (Kreibig, 2010). Alternatively, couplings between the subsystems might be absent in self-reports of emotion because of a subliminal stimulus, unconscious emotions, or emotion regulation (Kreibig, 2010; Öhman, Carlson, Lundqvist & Ingvar, 2007).

What does this multi-component and dimensional understanding of emotion imply for the measurement of emotions? Mauss and Robinson (2009) suggest that the lack of strong convergence among multiple measures of emotion implies that the construct of 'emotion' cannot be captured with any one measure considered in isolation. They conclude that the more measures of emotion that are obtained, and the better tailored they are to the particular context and research question, the more likely it is that a researcher will learn from a particular study. Since different measures of emotion appear to be sensitive to different dimensional aspects of an emotional state (with EDA sensitive to arousal, facial expression sensitive to valence, etc.), they can be expected to act in a complementary way.

To sum up, a multiple-component, dimensional understanding of emotions (with not much coherence between the response systems) implies first of all that we need to measure all three response systems. Thus, we need to use multiple methods in such way that they cover all three subsystems (subjective experience,

behaviour, and the ANS). Secondly, the dimensional understanding implies that we need to focus on the dimensions of both intensity and valence. Thirdly, our interest in investigating emotions in connection with learning processes implies that we need to measure how emotions elicit learning processes and outcomes – and vice versa, i.e. how learning processes elicit emotions.

The following section addresses the strengths and limitations of different measures as they encompass subjective experience, the ANS, and behavioural subsystems of emotion.

3. The strengths and limitations of self-reporting, physiological, and behavioural indicators of emotions in learning

The self-reporting of emotions has several clear strengths. Pekrun (2016) suggests that self-reports can render more differentiated assessments of emotions than any other current method. Self-reports are especially important for a nuanced description of emotions and thoughts. From a practical perspective, self-reporting is a more economical means of data collection than (for example) observations (Pekrun, 2016).

However, self-reports also have clear disadvantages, since they are limited to those emotional responses that a participant is aware of, i.e. to emotions that are inside her/his conscious mind. Thus, they do not capture unconscious emotional processes (for unconscious emotions see Winkelman & Berridge, 2004). In addition, if self-reports are collected retrospectively, memory-bias is common. Kreibig (2010) suggests that self-reports of emotions are likely to be more valid to the extent that they relate to currently experienced emotions. Even in this case, however, there are concerns that not all individuals are aware of and/or capable of reporting on their momentary emotional states (Mauss & Robertson, 2009).

In addition, there are differences between cultures in the understanding of terms relating to emotions, as well as between individuals sharing the same cultural background. An additional threat to validity emerges from the human tendency towards impression-management, which is always present in the social context of data collection (Azevedo et al., 2016). The bias pertaining to impression-management and the displaying of emotions might actually be particularly strong among professionals such as teachers and leaders, whose professional work competences comprise the skills of displaying and regulating their own emotions (e.g. Kreibig, 2010). Nevertheless, even if it were easy to change how one reports one's emotions for the purposes of impression-management, it may be much more difficult to change one's level of physical activation in efforts at coping and impression-management (Azevedo et al., 2016). From this point of view, physiological and behavioural indicators are needed to increase validity and reliability in the measuring of emotions.

All this would imply that self-report methods should not be used alone, and should be supplemented with psychophysiological and behavioural indicators. This would also accord with an understanding of emotions as taking place in three subsystems which do not cohere, such that the measures of one cannot substitute for another. Azevedo et al. (2013, 2016) have suggested the use of *electrodermal activity* (EDA), which has long been used as a marker of sympathetic nervous system activity. In addition, researchers addressing emotions in connection with learning have suggested the use of facial expressions and eye-tracking to detect, identify, and classify affective states during learning (Azevedo et al., 2013, 2016). EDA refers to changes in the skin's ability to conduct electricity. It can be measured by well-established and non-intrusive techniques that provide unique information on emotional arousal, increased cognitive work load, and task-engagement. Overall, the higher the conductance level rises, the more elevated the subject's emotional arousal becomes (Huhtamäki et al., 2017; Laitila et al., 2018). It has been suggested that if EDA is a good predictor of cognitive load, task difficulty, and task engagement, it might also be used for predicting boredom, disengagement, and frustration in learning (Azevedo et al., 2013, 2016).

A feature of EDA is that it is highly sensitive to small changes in low states of anxiety (Boucsein, 2012; Hugdahl, 1995). In addition, EDA is responsive to behavioural inhibition and to defensive strategies. Thus, EDA increases if thoughts are suppressed or if emotional expressions are inhibited. Nevertheless, an increase in EDA is not equivalent to the occurrence of emotions with negative valence or stress; in fact, EDA

is also responsive to increased activation within the body in the case of happy excitement (Benedeck & Kaernback, 2010; Karvonen, 2017).

The limitations involved in using EDA are often positioned around the difficulty of extracting information from the 'raw' EDA signal. In addition, there can be significant daily variation in participants' EDA response, with EDA levels increasing linearly throughout the day, and even from one season to another. From such considerations, Azevedo et al. (2013, 2016) recommend including EDA, but not limiting oneself to it, in efforts to measure ANS activity.

A central technique for studying ANS emotion responses has been *heart rate variability* (HRV), which refers to a variety of methods for assessing the beat-to-beat change in the heart over time (Quintana & Heathers, 2014). HRV describes the variation between consecutive inter-beat intervals. Both sympathetic and parasympathetic branches of the ANS are involved in the regulation of heart rate (HR). Sympathetic nervous system (SNS) activity increases HR and decreases HRV, whereas parasympathetic nervous system (PNS) activity decreases HR and increases HRV (Brentson at al., 1997; Tarvainen et al., 2018).

HRV is commonly used as a tool in assessing cardiac autonomic regulation, and it has been used in a range of wellbeing applications to evaluate the functioning and balance of the ANS (Tarvainen et al., 2018). It has been suggested that HRV can be used to investigate the relationships between autonomic regulation and interpersonal interaction (Porges, 2001). Increased HRV has been found to indicate a feeling of safety. In group-based learning contexts, such a feeling has been connected to having space for self-reflection and the emergence of new ideas. In contrast, reduced HRV has been observed in disorders characterized by poor social cognition and emotion regulation (Quintana & Heathers, 2014).

Despite this, there are severe limitations concerning the use of HRV. Because of the large inter-individual variation in HRV, it has been suggested as more appropriate to a within-subject design. In addition, HRV is affected by respiratory depth and frequency, with both breathing and blood pressure regulation having their own directly mediated relationships with HRV. Thus, social-emotional tasks that induce changes in respiratory time variables and/or depth may indirectly influence HRV. In addition, a number of studies have shown continuous focused attention (and thus tasks with increased attentional demands) to reduce HRV, primarily because of changes in respiratory depth and frequency. This creates further difficulties for interpretation (Quintana & Heather, 2014; Quintana, Alvares & Heathers, 2016). The starting point for HRV analysis is the electrocardiogram (ECG) recording, from which the HRV time series can be extracted. In the formulation of HRV time series, the fundamental issue is determination of the heart beat period (Tarvainen et al., 2018).

In addition to the ANS measures, researchers addressing emotions connected to learning have suggested the use of *facial expressions* and *eye-tracking*. These might allow researchers to detect, identify, and classify affective states during learning (Azevedo et al., 2013, 2016). In fact, the use of facial expression is a fundamental process-oriented approach in the detection and classification of affective states (Azevedo et al., 2016; Ekman, 1992). In everyday social interaction, facial expression has a clear communicative function, but it also functions as an automatic and observable indicator of expressed emotion. Recently, data collection on facial expression has been conducted through software linked to a video data stream of the learner's facial expressions. In addition, there are now many comprehensive, widely supported methods to objectively describe the facial expressions of emotions, based on Ekman's Facial Action Coding System (FACS). These analyse basic emotions described as (for example) *enjoyment, fear, anger, sadness, disgust*, and they indicate how the emotions evolve over time. Commercial applications for facial expression recognition software have been developed to automate the coding process involved.

Despite the use of automated software to register the duration, fluctuation, transition, and dynamics of affective states, Azevedo et al. (2016) emphasize that several disadvantages should be considered. First of all, the systems rely heavily on the quality of the video stream. There must not be shadows on the faces. Moreover, eyeglasses can make it difficult to measure some emotions, and can increase the likelihood of interpretation as a neutral gesture. In addition, postures and body movement are limited, the software can only identify one face at a time, and the results are restricted to sets of specific predefined emotions. Despite this, Azevedo et al. (2013; 2016) suggest that the benefits of automatic facial expression recognition software outweigh the

limitations. It permits non-intrusive, reasonably valid, and concurrent measurements, which can be integrated with other process data channels. Thus, the method has clear potential significance.

Eye tracking is commonly used in learning research. The data can provide detailed description of the areas and focus of the subject's interest (Azevedo et al., 2016; Hautala, Loberg, Hietanen, Nummenmaa & Astikainen, 2016). Eye-tracking data can be used to measure the learner's fixations (Where?), fixation durations (How long?), saccades (eye movements from one fixation to another), plus the order and the gaze patterns (multiple fixations) that occur during learning. From these data, we can gain insights into what learners pay visual attention to during learning, and this can be both indicative and predictive of emotions (Bondareva, Conati & Feyzi-Behnagh, 2013). Nevertheless, despite the strengths of eye-tracking data for research on emotion and learning, eye tracking alone cannot definitely indicate what elicited the emotion; nor can one determine what the effect of this emotion may be on subsequent activity. Thus, eye-tracking data must be supplemented by other data collection methods if one is to capture the subject's sense-making processes, as they take place in the reciprocal relationship between learning and emotions.

The *Stimulated Recall Interview* (SRI) method has been increasingly used as the video-recording of learning situations has become more popular (Kagan, Krathwohl & Miller, 1963). Especially in group learning contexts, when one can organize subsequent viewing of the group situation and of one's ways of acting in the situation, there is the cognitive capacity and space to gain new insights, via reflection and re-evaluation of one's own actions (Huhtamäki et al., 2017; Lyle, 2003; Vall et al., 2018). In teacher education, the video-assisted procedure has long been used to increase self-reflection, and thus possibilities to learn about oneself as a teacher (Fuller & Manning, 1973). These goals (involving an increase in self-reflection, and an understanding of oneself as a leader) were in fact the learning goals in the Leadership Coaching Program addressed in this study. Hence, SRI was expected to be a useful method for our purposes. We expected that viewing the videotaped learning situations in connection with the perceived emotions would provide space for reflection on the original learning situation (at a first viewing) with insights into the emotions experienced (at a second viewing).

4. Research questions and aims of the study

The research questions can be specified as follows:

1. What kinds of information do concurrent self-reports, indicators of ANS, and behavioural measures of emotion provide, in terms of understanding a person's emotional responses within professional learning?

2. To what extent does the emotion-driven Stimulated Recall Interview (SRI) promote reflection and hence learning, related to the original learning situation?

The prime aim of the study was to form tentative conclusions on how self-reports, indicators of ANS, behavioural measures of emotion, and the emotion-driven Stimulated Recall Interview (SRI) may provide complementary information on emotions in professional learning. A further aim was to elaborate the potentials and challenges of a multi-componential methodology for researching emotions in professional learning.

5. Methods

5.1. Design, procedure, and stimulus material

The multi-method measurement procedure described below was designed for ultimate use in wider data collection. In the present case we applied the procedure and data collection setting to a single subject. A

practical purpose in this pilot study was thus to elaborate what might need to be changed or developed for wider data collection.

The ethical committee of Jyväskylä University evaluated the design and approved the study. All the participants mentioned in the article, and all participants at other times, gave their informed consent regarding participation.

Figure 2 gives a general description of the procedure (including the stimulus material and data collection methods) used in this study. Figure 2 shows that in session I, the subject viewed four successive video-recorded episodes (A1, B1, B2, A2) and assessed these via the on-line *Emotion Circle* tool, described in detail in Section 5.2.1. In session II, which followed immediately after session I, an emotion-driven Stimulated Recall Interview (SRI) was conducted while the subject again viewed the same four episodes (A1, B1, B2, A2), which were shown together with the saved EC data from her emotion assessments in session I. Psychophysiological data were collected via measurement of EDA, HRV, and respiration. Electrodes measuring EDA were attached to the left hand, HR electrodes to the chest, and the respiration belt around the chest (see Figure 2). Behavioural data from eye movement were gathered via the BeGaze Version 3.7 (SMI) recorder, which was located at the lower edge of the display. The Face Reader 6.1. (Noldus) video camera (which was used for recording the subject's facial expressions) was located behind the screen (see Figure 2).

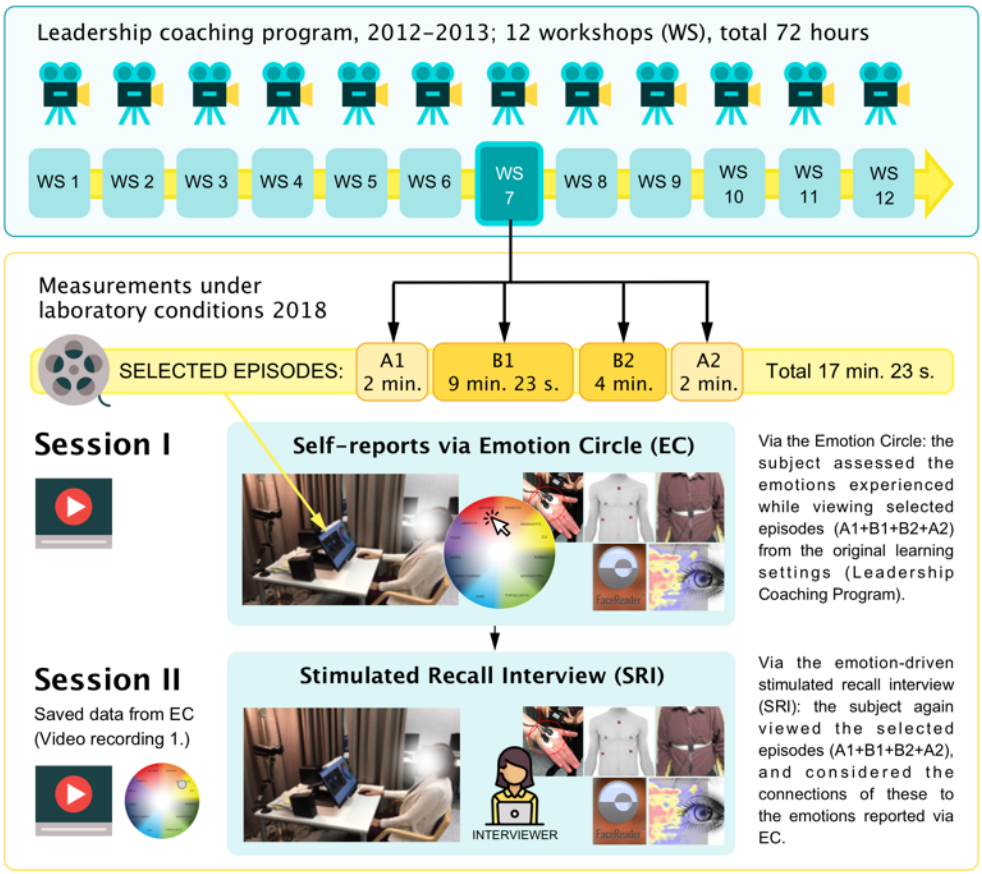

Figure 2. Procedures used in this study.

Sessions I and II took place under laboratory conditions in April 2018. In contrast, the selected episodes (A1, B1, B2, A2), which were used in the laboratory settings as stimulus material, were selected from the *original learning settings.* These had taken place five years earlier (in 2012–2013.) Our subject, Lisa, had participated in this earlier group-based training platform (constituting a Leadership Coaching Program). The

program was constructed to cultivate (i) participants' professional identities, (ii) their ways of managing social relationships within the work community, and (iii) their professional communication (Hökkä, Vähäsantanen, Paloniemi & Eteläpelto, 2017; Vähäsantanen, Paloniemi, Hökkä & Eteläpelto, 2017).

The program had twelve workshops in all. It covered a period of eleven months, with one day per month allotted to it. The selected episodes B1 and B2 were those which Lisa had mentioned as representing the most important and meaningful learning situations for her, personally, within the program. Episodes A1 and A2 were selected as representing neutral episodes. These were episodes which Lisa had not mentioned at all. All of them came from the seventh workshop. The procedure of selecting and ordering the episodes to be viewed was constructed according to the ABBA model. This meant that at the beginning and at the end of the session there were emotionally neutral episodes. The episodes Lisa had selected as most influential in terms of her learning (B1 and B2) were placed in the middle of the temporal continuum. The episodes differed in length, being chosen in such a way as to comprise coherent authentic learning episodes, bearing in mind that they would otherwise have been too fragmented for the subject to understand and evaluate. Under laboratory condition, the four episodes were played consecutively (total duration = 17 min. 23 s.). The contents of the episodes were as follows:

Episode A1 (2 min.) depicted an informal get-together situation before the formal start of the training. In this episode, group members came to the seminar room in which the group learning session would take place. About half of the participants, plus the trainer, were already standing there face to face. They were chatting and talking informally with each other. One by one, the remaining participants walked into the room. Some of them went directly to sit down in chairs placed in the form of a half circle. There was mix of voices, thus it was impossible to differentiate individual voices, although one could see the people who entered the room.

Episode B1 (9 min. 23 s.) depicted a situation in which some of the participants (who were close to each other as colleagues) talked about their present feelings using *symbolic object working*. This took place in such a way that the participants, plus the trainer, (13 persons altogether) were sitting in chairs which were in the form of a circle. In the middle of the circle there was an empty chair. This symbolized an empty place in which each participant could place an imagined object, one that best represented a central issue in their life at that moment. At the start of episode (B1), a participant, 'Bertha' started to describe how busy she was in her work at that moment, and how she had tried to organize some weekend trip in order to recover a little; however this had merely caused more stress and internal conflict in her mind. The trainer put some questions to Bertha concerning the issues that would emerge if there was space for something other than work. After this, another person ('Caroline', a close colleague of the previous speaker) started to talk. She described very similar feelings of having too much work, with very stressful feelings connected, for example, to the salary negotiations taking place in her work organization. Caroline continued to talk about serious issues regarding her health and wellbeing. While talking about these stress-induced health issues, and how they were connected also to family issues, Caroline burst into tears. After this she expressed surprise about the strong emotions connected to her situation. Immediately after the start of the weeping, Lisa stood up. She picked up a tissue from a side table and gave it to the crying person. One by one, two other close colleagues of Caroline started to cry (indicating emotional contagion). The remaining participants looked very serious. The trainer put further questions concerning how far Caroline had listened to her own mind, and how she could take more time to take care of her own health and wellbeing. The situation ended with the talk of a third person, who was also crying. At the end of the episode, that person made a joke. This caused people to laugh, and feel more relieved.

Episode B 2 (4 min.) depicted a situation exemplifying how to work with difficult cases. The work had started with a preliminary task in which each participant had called to mind, from their personal work history a difficult colleague. They had described the situation in writing before the session and sent it to the trainer. In this episode, participants first sat face to face, in the form of a circle. The subject of the present study, 'Lisa', role-played her difficult case using a drama method. She took the role of the difficult person, changing her voice and way of talking, to resemble that of the difficult person. After this role-play, in the next part of the episode, the participants turned their chairs 180 degrees, so that they were no longer face to face. From this 'turning chair' position, Lisa presented what she was actually thinking, and what she would have liked to say (as her authentic self) to the difficult person. She was then saying aloud what she would have

wanted to say to the imagined difficult person, and what she actually could not have said in real life. After this, the chairs were once again turned to the face-to-face position. Lisa now explained to the whole group the kind of history she has had with the difficult case. At the beginning of each piece, the trainer always put a question to be answered. In presenting her reasons regarding the difficulty of the person, Lisa also described how she had at last become empowered to set limits to the negative and destructive behaviour of the difficult person.

Episode A2 (2 min.) depicted a pair-discussion session of the whole group. The participants were sitting in chairs, which were in the form of a circle. They actively discussed with the person next to them. There was mix of voices; hence, one could not differentiate individual voices. However, one could see smiling faces, loud laughter, and the participants' active concentration during discussion with their pairs.

5.2. Data collection

5.2.1. Self-reports via the Emotion Circle (EC)

Self-reports were collected from Lisa via the *Emotion Circle* (EC) on-line assessment tool. This was developed for concurrent assessment of the quality and intensity of emotions. Here it should be noted that there has been a lack of valid, user-friendly tools to capture changing emotions during learning processes. In our research project, we developed the EC on-line application for the self-assessment and reporting of individual shifting emotions within professional learning settings (Eteläpelto et al., 2017). EC utilizes a colourful graphic interface containing 12 written emotion words (Figure 3). These are presented, in line with the *circumplex* model of emotions (Russell, 2005). The circumplex model has been designed as a single-item scale providing a quick means of assessing affect along the dimensions of pleasure – displeasure (valence) and arousal – sleepiness. Pleasure is considered to be the bipolar opposite of displeasure, and the subjective feeling of arousal to be the bipolar opposite of sleepiness. These dimensions are further considered to be orthogonal (i.e. independent) and thus conceptually separated (Russell, Weiss & Mendelsohn, 1989). The single-item scale based on these two dimensions is envisaged as an instrument that will be short and easy to complete in assessing the subjective experience of continuously and rapidly fluctuating emotions in repeated-measures design (Russell, Weiss, & Mendelsohn, 1989). One can anticipate that multiple-item checklists or questionnaires would be too time-consuming and distracting for the purpose of reporting continuously changing emotions within the learning process. In the circumplex model of emotions, valence is described along the X axis so that on the left there are unpleasant (negative) and on the right pleasant (positive) emotions. Arousal is described along the Y axis so that on the upper segment (containing plus values) there are emotion constructs characterized by high arousal ('hot' emotions) whereas on the lower segment (containing minus values) there are emotion constructs characterized by low arousal ('cold' emotions).

In accordance with the circumplex model of emotions, different emotion words are placed in the Emotion Circle (EC) (see Figure 3). This means that on the upper right quadrant of the circle, there are emotions characterized by pleasant activation and activated pleasure, such as excitement, surprise, and joy. On the lower right quadrant there are emotions characterized by deactivated pleasure and pleasant deactivation (safety, compassion, courage). On the upper left quadrant there are emotions characterized by unpleasant activation and by activated displeasure (irritation, anxiety, and fear). On the lower left quadrant there are emotions characterized by deactivated displeasure and by unpleasant deactivation (shame, frustration, and sadness).

The intensity of emotions is depicted in the EC via degrees of colour saturation. In the middle of the EC the colours are lighter, denoting less intensity of emotion. As one moves to the circumference, the colours become more saturated, denoting more intense emotions.

Figure 3. The display of the Emotion Circle (EC).

Figure 3 shows the EC display. The emotion constructs used in the Emotion Circle were selected on the basis of interviews conducted with the 11 participants of the Leadership Coaching Program. Within the interviews, the interviewees were first presented with an open question concerning their perceived emotions during the program. They were then shown a list of 28 emotion constructs that were expected to represent the most common emotions felt during the program. These emotions were based on the final interviews conducted at the end of the program in 2013 (Hökkä, Vähäsantanen, Paloniemi & Eteläpelto, 2017). In selecting the relevant emotion constructs, we also utilized prior studies on emotions in leaders' identity negotiation (e.g. Winkler, 2018).

Self-reports have been criticized on the grounds of the subjects' tendency to report more positive emotions because of a social preference for these. Because of this, EC included roughly the same number of positive (EC right side) and negative (EC left side) emotionally-related words. It was anticipated that this would help to counteract any tendency towards the reporting of positive rather than negative emotions.

EC automatically saves the process data (time and object of clicking) from the subject's assessments. This assessment video (in the present case Video recording 1) can be shown to the subject immediately after the assessment.

The EC application makes it possible to collect subjects' self-assessments of their situation-specific emotions, including also data on the quality and intensity of the emotions, plus their dynamic continuity. It transforms and displays the process in such a way that it can be synchronized with other (physiological and behavioural) process data, collected at the same time. The EC also makes it possible to show the subjects' recorded assessments together with the video-recording of the situation which the subjects have assessed. This is needed for stimulated recall. In using EC, we seek to avoid the memory bias of retrospective interviews.

When we used the EC, the subject, Lisa, was first given general instructions on it, with opportunities also to practise the use of it. She was asked to click on the emotion word or words which in each situation represented her subjective experience of the emotion. In order to guarantee that she had properly understood the use of EC, she was asked to imagine some emotion, then to click on the EC accordingly. After Lisa had confirmed that the use of EC was easy for her, we played the recorded episodes in the order A1, B1, B2, A2, using EC to collect her self-reports, including the nature and intensity of her emotions.

5.2.2. Behavioural data

We obtained behavioural and expressive data from Lisa using automatic gaze recordings (via BeGaze) and video-recording of gestures (via Noldus Face Reader). The methods used are unobtrusive. They include a gaze recorder, located at the lower edge of the display. A video camera is used to collect data on facial expression (Face Reader data). This is located behind the display, with an additional light placed on the upper side of the display to prevent shadows on the faces.

5.2.3. Autonomic nervous system recordings

Autonomic nervous system (ANS) recordings were taken from Lisa during the video viewing and assessment situation. This further continued during the stimulated recall session which followed the viewing and assessment. During the ANS recording sessions the following signals were recorded using the QuickAmp amplifier and data acquisition system (Brain Products, Gilching, Germany):

The electrocardiogram (ECG) was recorded with two Ag/AgCI electrodes (Ambu Neuroline 710, Ballerup, Denmark), attached above and below the heart, with a similar ground electrode attached over the stomach.

Electrodermal activity (EDA) during the session was recorded via two skin conductance (SC) electrodes (EL507, Biopac Systems, California, USA) on Lisa's non-dominant palm, below the first and fourth digits. The palm was chosen as the location, because in piloting and in previous research (Karvonen, 2017), it was found that there is less measurement error from hand movements when the electrodes are in that area, as compared to fingertips. SC was determined using 0.5 V constant voltage (GSR sensor, Brain Products, Gilching, Germany).

Respiration during the session was registered via a fabric belt (Respiratory effort sensor, Spes Medica, Genoa, Italy). This was fastened on top of Lisa's clothes, on the lower chest area. However, respiration was not analysed in this study, because the data quality was found to be inadequate (due to the belt having been too loosely attached, as discovered after the session).

EDA and respiration were amplified in DC mode, but the ECG was 0.5 Hz high-pass filtered. Signals were acquired with a sampling frequency of 1000 Hz, using a data acquisition program (BrainVision Recorder, Brain Products, Gilching, Germany).

A custom-made marker unit was used to synchronize the ANS measures to the video.

5.2.4. The Stimulated Recall Interview (SRI)

After Lisa's assessment of her emotions during viewing of the video-recorded learning episodes, a video- and EC-assisted *Stimulated Recall Interview* (see SRI; Kagan, Krathwohl & Miller, 1963) was conducted (immediately after the assessment). In this interview, the video episodes were shown to her, along with her emotion assessments given via EC. She was encouraged to share her thoughts, feelings, and reasons at any time while watching the videos, including her assessments of the emotions that she had experienced while watching the videos, and during the original situation as she recalled it. When she started to speak, the video was stopped to give her time to explain and describe her thoughts.

To assist in this phase, Lisa was given general instructions, in the form of questions to consider, as follows: *What thoughts, feelings, or bodily sensations did you have while watching the video and assessing your emotion connected to it?* We assumed that naturally, she could have forgotten many of the feelings she had experienced during the original learning sessions (which had taken place five years previously), and that she might simply describe the thoughts, feelings, and sensations evoked by watching the session videos.

Nevertheless, it seemed reasonable to suppose that she might be able to recollect some intense emotions from the original learning sessions. For this reason, she was further asked to specify *whether she had had a particular thought or feeling in the original learning session, or whether that thought or feeling emerged only now in the assessment session*.

The same ANS measures were recorded during the SRI as during the emotion assessment session, and the interview was recorded with a video camera. After the Stimulated Recall Interview, Lisa was further asked to comment on the user-friendliness of the Emotion Circle, and of the data collection sessions as a whole.

5.3. Analysis of the data

So far, there have not been many analytical (and especially statistical) techniques which would be relevant for analysing data from single-case research, characterized by the rapid and randomly determined alternation of conditions or processes. In their review of existing analytical techniques in single-case design, Manlov and Onghena (2017) found visual analysis, constituting the classical way of analysing single-case data, to be the most frequently applied technique. It has been suggested that in most cases, visual analysis is sufficient to demonstrate evidence of a relationship between conditions and outcome variables (Kratochwill et al., 2013). A baseline phase is necessary to represent the control condition and to provide a clear basis for comparison. Visual analysis allows comparisons between and within phases, indicating levels, trends, and variability, with data also on overlaps, the immediacy of the effect, and the consistency of patterns. Visual analysis can be used to suggest a functional relation between conditions and the target behaviour, and it can indicate the most salient features of the data. Visual data can be used as an initial step in the analysis, and it can be regarded as complementary to statistical analysis in making sense of the data obtained (Manlov & Onghena, 2017).

In this study we used visual analysis to compare the levels, trends, and variability of the ANS data on EDA and HRV, within and between the four time segments (episodes A1, B1, B2, and A2), referring also to the pause segment, which was used to give a baseline value for the comparisons. Visual analysis was also used to carry out a general assessment of the data patterns in EDA and HRV.

Statistical analysis was used as complementary to visual analysis in the analyses of EDA data. From the EDA data, statistically significant ($p<.05$) values were calculated.

In addition, visual analysis was applied in the descriptive analysis of the concurrent variation of different data modalities. This was done in respect of the ANS data and the on-line self-report data on emotions collected via the Emotion Circle. In the visual analysis addressing the complementarity of ANS data and self-reports, a figure was constructed using the time stamps written on the Excel file. This indicated the exact time points and the specific emotion assessed with EC while the subject reported her emotion. In this visual description, emotions with a positive valence (excitement, surprise, joy, courage, compassion, safety) were placed on the upper side of the horizontal time axis, while emotions with a negative valence (sadness, frustration, shame, fear, anxiety, and irritation) were placed under the time axis. Specific colours connected to the different emotions in EC were also presented in the visual description (Figure 5.).

From the EC data we calculated the absolute and percentual frequencies of clicks on different emotion words within each episode.

Data from the subject's *EDA* were analysed with the Ledalab program (version 3.4.6) written in Matlab (Benedek & Kaernbach, 2010; www.ledalab.de). Before the analysis, the sampling rate was reduced to 10 Hz, which was high enough to represent rapid changes in SC related to SNS activation. The rapid components of SC were extracted as skin conductance responses (SCRs) and written to an Excel file. The SCRs were normalized by computing the average and standard deviation of the session, and calculating z-scores. Values larger than 2.0 were considered to be statistically significant at the $p<.05$ level (given that 5% of the values have that property; they can thus be considered to represent statistically significant SNS activation).

The rationale behind the analysis of the statistically significant EDA peaks is based on the assumption that EDA can track rapid and unconscious changes in sympathetic nervous system (SNS) activity in very brief time windows. An increase in SNS activity is related to the increased physiological arousal that accompanies

most emotions and also preparation for action (Boucsein, 2012; Kreibig, 2010). In particular, rapid changes in EDA (measured as skin conductance responses, and indicated by increased sweating, especially in palms, fingers, and feet) are thought to be a direct measure of the phasic neuronal activity of the SNS (Benedek & Kaernbach, 2010). Laitila et al. (2018) have demonstrated the added value gained from analysing the EDA responses that occur in the social interaction of a couple therapy session, as a means to detect important moments of change at individual and interpersonal level. In our subjectively meaningful social learning sessions, we were also interested in detecting critical moments of change that had made the selected episodes B1 and B2 personally meaningful for Lisa's learning within the Leadership Coaching Program. In line with the analysis of Laitila et al. (2018), we expected that those values of EDA which represented rare ($p>.05$) high peaks would reveal exceptionally high unconscious emotional arousal of the subject, and thus point to critical moments of learning. In the analyses of Lisa's EDA, we compared the numbers of statistically significant high peaks between episodes of about the same length. The baseline value during the pause was used in the comparison.

Nevertheless, more important than mere detection of moments of change is determination of what has led to these moments, thus placing the focus on the learning process itself (Laitila et al., 2018). This implies that if one is to make sense of EDA data in terms of learning, they need to be complemented with other kinds of data. In the present study, ANS data from EDA (indicating rapid peaks in the activity of the SNS) were complemented with HRV data, indicating the activity of the parasympathetic nervous system (PNS). In contrast with the SNS responses manifested in EDA, the PNS responds much more slowly, possibly over minutes rather than seconds. Thus, during fairly short episodes (from 2 min to 9 minutes) there could be instances of overlapping from one episode to another. This is a feature that needs to be considered in interpreting HRV trends and changes.

HRV was analysed in the time domain using the Kubios HRV Premium program (version 3.1; www.kubios.com). First of all, R peaks were detected in the ECG to determine the intervals between consecutive heart beats. Possible artefacts were removed automatically. RR intervals were determined, and root mean squares of successive RR interval differences were determined for 60-second windows, starting from the beginning of the session, and covering the whole session in 10-second steps. The HRV was not normalized (in order to keep the relevant values transparent in milliseconds).

Face Reader is based on the circumplex model of emotion. The model describes emotions in a two-dimensional circular space, containing arousal on the vertical axis and valence on the horizontal axis. The centre of the circle represents a neutral valence and a medium level of arousal. The circumplex model of FaceReader Version 6.1. is based on the model described by Russell (1980). The valence in Face Reader indicates whether the emotional status of the subject is positive or negative. 'Happiness' is the only positive emotion, while 'Sadness', 'Anger', 'Fear', and 'Disgust' are considered to be negative emotions. 'Surprise' can be either positive or negative. The valence is calculated as the intensity of 'Happiness' minus the intensity of the negative emotion with the highest intensity. For instance, if the intensity of 'Happiness' is 0.8 and the intensities of 'Sadness', 'Anger', 'Fear', and 'Disgust' are 0.1, 0.0, 0.05, and 0.05 respectively, then the valence is 0.7 (see Face Reader Version 6.1. Reference Manual, 2015, pp. 80–81).

The focus of visual attention was obtained via beGaze, together with the subject's assessments conducted with the on-line Emotion Circle. These process data were analysed using a process analysis, conducted via the video recordings, as depicted in the attached video depicting episode A2 (see Video 1). The process analysis derived from the video recordings demonstrates how the focus of visual attention always preceded the subject's selection of the emotion words she clicked. This information can be used in further analyses of the subject's selection process, as well as in analyses of the usability of the Emotion Circle. In this study, the data were used to demonstrate the complementarity of the two relevant data modalities (comprising the more objective behavioural gaze data, and the subjective self-assessment data derived via the Emotion Circle).

A2 EC assessment.mp4

Video 1. Assessment process (2 min. 3 s.) as depicted for the process analysis with the assessments (red square) conducted via the Emotion Circle, and via the focus of gaze (blue circle). Data from episode A2 (MP 4 format file).

For the purposes of detailed process analysis, data derived from different methods can be further transformed via the Open Broadcaster Software, changing the data into the MP4 format. This format can be moved to The Observer XT12 for simultaneous display. In this way different data sets can be scrutinized simultaneously, and observed moment by moment on the screen.

The SRI interview data were transcribed verbatim, amounting to 3.5 pages (A4, 1.5 line space). The transcribed data were read and re-read by the second and first authors. We first identified verbal expressions of emotions, plus the explanations given for these. Secondly, we identified self-reflective or other reflective contents of the utterances. Finally, we focused on displays of insight and novel ideas which were not mentioned in the video-recorded episodes.

6. Findings

The findings are presented here in line with the research questions. We first describe the kinds of information available from concurrent self-reports, obtained via the Emotion Circle, the ANS indicators, and the behavioural measures of emotion. These shed light on the components of emotions and learning processes, and the emotional responses that occur in professional learning (6.1). Thereafter (6.2) we present findings based on the emotion-driven Stimulated Recall Interview (SRI), with comments on how these were connected to the learning situations in question.

6.1. Information elicited via concurrent self-reports, ANS indicators, and behaviour

Here we first describe findings from Lisa's self-reports on emotions, derived via the on-line Emotion Circle (6.1.1). We then address the facial expressions obtained via the automatic face reader; these functioned as complementary data to verify the self-reported emotions (6.1.2). ANS data concerning electrodermal activity (EDA), and Heart Rate Variability (HRV), are presented in 6.1.3.

6.1.1. Self-reporting of emotions via the on-line Emotion Circle (EC)

The analysis of Lisa's self-reporting of emotions via the on-line Emotion Circle (done while watching the selected episodes of videotaped learning sessions) showed that during the four episodes (17 min. 23 s. in total), she clicked on emotions 134 times. Table 1 shows that the clicks covered the whole time period fairly evenly. As expected, the quality of the emotions was different between the four episodes. Emotionally neutral episodes (A1 and A2) were placed at the beginning and end of the session, with B1 and B2 being placed centrally. These were the episodes arousing the strongest emotions, and also the episodes which Lisa had selected as most influential for her learning.

Regarding the contents of the emotions reported via EC, Table 1 shows that only one of the given emotion words (*irritation*) was not reported at all. The most used emotion word was *compassion*. Lisa would have liked to add to the EC the experience of feeling *guilt*, especially in episode B1. Table 1 shows that emotion words with positive valence (used 100 times) were used three times as much as those with negative valence (used 34 times).

Table 1 further shows that the quality of Lisa's emotions was quite different between the four episodes. Episodes A1 and A2 were fairly positive overall, since all the reported emotions exhibited positive valence (characterized by intrinsic pleasantness). As opposed to this, episode B1 included strong negative and

unpleasant emotions, such as *anxiety*, *sadness*, and *fear*. These were accompanied by a high degree of compassion. By contrast, episode B2 was assessed as fairly positive in terms of emotions such as *courage, surprise, joy*, and *excitement*. In addition, episode B2 was reported as displaying *safety* and as including also *compassion*.

Table 1
Number (absolute frequencies and percentages) of emotion words used in Lisa's assessment of four successive video-recorded episodes (A1, B1, B2, A2) via the on-line Emotion Circle (EC)

Emotion word	Episode A1 2 min.	Episode B1 9 min. 23 s.	Episode B2 4 min.	Episode A2 2 min.	Sum 17 min. 23 s.
Excitement	1 (6%)	1 (1%)	3 (10%)	2 (14%)	7 (5%)
Surprise	12 (75%)	2 (3%)	7 (24%)	2 (14%)	23 (17%)
Joy	3 (19%)		4 (13%)	5 (36%)	12 (9%)
Courage		1 (1%)	9 (30%)		10 (8%)
Compassion		36 (49%)	4 (13%)	2 (14%)	42 (31%)
Safety		1 (1%)	2 (7%)	3 (22%)	6 (5%)
Sadness		8 (11%)			8 (6%)
Frustration		5 (7%)			5 (4%)
Shame		3 (4%)	1 (3%)		4 (3%)
Fear		7 (9%)			7 (5%)
Anxiety		10 (14%)			10 (7%)
Irritation					
Sum	16 (100%)	74 (100%)	30 (100%)	14 (100%)	134 (100%)

While watching the videos, the perceived emotions were assessed by clicking on the emotion words given in the Emotion Circle (Session I). However, this is a somewhat limited way of expressing one's emotional experience. Nor, taken by itself, does it indicate or give information on the *reasons* for the specific reported emotions. Because of this, the EC assessments were further elaborated (in session II) using the Stimulated Recall Interview (SRI) method (6.2).

6.1.2. Face reader and valence of emotions

Differences between the episodes in terms of the valence of emotions were further analysed with the Face Reader (see Figure 4). It confirmed the findings based on Lisa's self-reports via EC. Episode B2 (Figure 4 right) was characterized by positive valence, whereas episode B1 (Figure 4 left) was full of very low and even negative valence. It should be noted that the glasses worn by the subject might influence error in counting gestures; thus, the findings showed a large amount of neutral emotion. Empty spaces in the graph indicate the Face Reader software not being able to identify Lisa's gestures because she was temporarily facing away from the video.

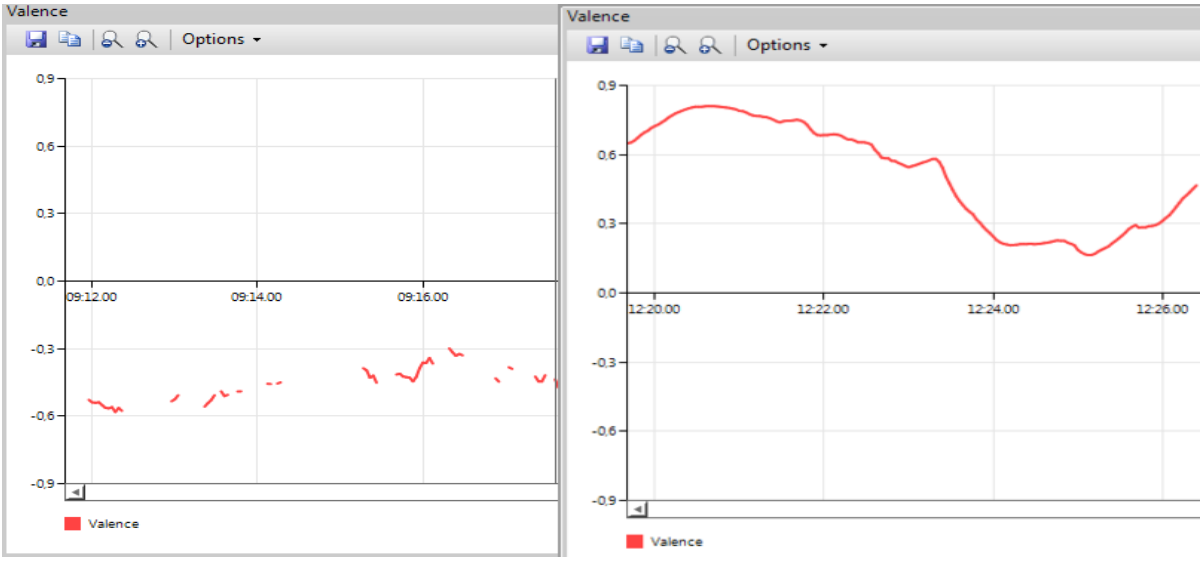

Figure 4. Screen captures of the valence of emotions within episodes B1 (left) and B2 (right).

6.1.3. ANS responses: EDA and HRV

Figure 5 shows the standardized EDA values and HRV values calculated for the four subsequent episodes (A1, B1, B2, A2) in Session I. In addition, in the middle of episode B1, there was a pause. This was not planned, and was due to technical problems. In fact, the sound of the video suddenly disappeared six minutes from the start of the viewing of the video, and four minutes from the start of episode B1, during the assessment of Lisa's emotions via EC. The technician then tried to work out the reason for the problem and make technical adjustments. This took eight minutes. During this time Lisa could do nothing but sit and wait for the adjustments to be completed.

In our analysis of the ANS data, we realized that this technical failure provided valuable baseline data concerning Lisa's EDA and HRV measures. At the start of the pause there was a rapid decrease in EDA. This remained low over the next six minutes. Just before the end of the pause there were some new attempts to recover the sound, with confusion and discussion concerning the functioning of the system. Lisa participated in these discussions, as indicated by some increase in EDA at the start of the pause. However, as Figure 5 shows, during the pause, her EDA values were at the lowest level for the entire data gathered during the session. The low EDA values during the pause could be connected to her passive (non-agentic) motivational state (being unable to do anything). This is in accordance with Kreibig's (2010) suggestion concerning low EDA: that it is characteristic of a passive state of action, and thus with deactivation of the sympathetic nervous system.

If the EDA levels of the pause situation are compared to the EDA levels during episodes A1 and A2 (which were selected to represent emotionally neutral episodes) one can detect a clear difference between the pause and the 'neutral' episodes. Especially in episode A1 (which came at the beginning of session I), the EDA levels were relatively high, with two statistically significant ($p<.05$) peaks within this episode. These peaks apparently manifested Lisa's pleasure at seeing other participants in the video, coming into the room in which the training had taken place. Episode A1 came at the start of the task of viewing and assessing emotions while viewing. Hence, this represented a novel way of working, one that might set an additional load (manifested as an increased EDA level). This is evident if one compares the EDA levels between episodes A1 and A2. The latter was at the end of assessment session I, i.e. at a point when Lisa was already familiar with the task. In the case of episode A2, she also knew that this was the last part of session I. She might therefore have felt more relieved than at the start of the session.

Nevertheless, if we focus (Figure 5) on the EDA levels during the personally meaningful episodes B1 and B2, we can see that there were many statistically significant high peaks in B1 (actually two before the pause, and 16 after the pause). In addition, the frequency of these peaks was fairly dense, and the frequency increased in the course of the episode B1. As described above, in Lisa's self-report concerning her emotions, episode B1 was described mostly in words with negative valence, such as *anxiety*, *sadness,* and *fear*. Here, she reported a high level of compassion with the person who was talking about her health problems. In the SRI, Lisa said that she felt as if she were the body of that person. Thus, the high EDA values here appear to be connected with the subjective experience of high negative stress (Kreibig, 2010).

Nevertheless, the high EDA levels were not connected merely to stress with negative valence. Figure 5 further shows that high EDA peaks were present also in episode B2, which was subjectively assessed as having fairly positive valence (i.e. in terms of self-reports). Episode B2 was described as including the subjective experience of *courage, surprise, joy,* and *excitement*. Lisa also reported feelings of *safety* and *compassion* (see Figure 5). This indicates that high EDA is not connected merely to emotions with negative valence, and can be related also to emotions with positive valence.

In general, Lisa's EDA levels here seemed to be related to active cognitive and affective work while viewing episodes that were personally highly important to her, and to assessing her emotions while watching these. These situations, depicted in episodes B1 and B2, were also those which she had selected as highly meaningful for herself in terms of her perceived learning outcomes within the Leadership Coaching Program. The EDA peaks, which were connected to watching these episodes, would thus indicate critical incidents in the learning processes, but also a high intensity of emotion. This is in accordance with prior research on EDA and its connections with the level of arousal, as manifested within active learning processes, and also in states of intense emotion (Kreibig, 2010).

Figure 5 also shows the values and changes in HRV in the course of the initial assessment session (I). As compared to EDA, which mainly indicates the activity of the more rapid sympathetic nervous system, HRV indicates the (generally more slowly fluctuating) activity of the parasympathetic nervous system. HRV was in the present case calculated in 60 s. time windows, which were moved forward in 10-second steps to cover the whole session. The calculation of HRV as the root mean square of successive heart beat differences (RMSSD) is not usually performed for periods shorter than 60 seconds (usually including 60–100 beats). Hence, computationally, the changes in the HRV could not be as rapid as in EDA, in which ten samples were used to represent phasic skin conductance responses in a second. An increase in HRV is a general marker of relaxing (Stein, Bosner, Kleiger & Conger, 1994). Nevertheless, there is also some natural variation in the HRV, and it is thus not directly connected to the events that take place in a given situation.

Figure 5 shows that there was slowly changing variation in the HRV within session I. At the beginning of the session, during episode A1, the HRV was very high, thus functioning as a marker of a relaxed body. However, at the end of episode A1, the HRV dramatically decreased. It increased again at the start of the next episode, B1. Thereafter, the HRV again decreased during the two EDA peaks before the pause. At the beginning of the pause, when there was confusion arising from the loss of the sound for technical reasons, the HRV again steeply decreased. After the start of this unexpected technical failure, the HRV then increased during the next six minutes of the pause time. This level of HRV provided Lisa's individual baseline in the passive situation, in which deactivation of the sympathetic nervous system was indicated by low EDA levels. After the pause, when episode B1 (characterized by negative valence) continued, the HRV tended to decrease throughout the episode, although the decrease was not linear. Especially at the end of episode B1, when there were many high and dense EDA peaks, the HRV was clearly decreasing.

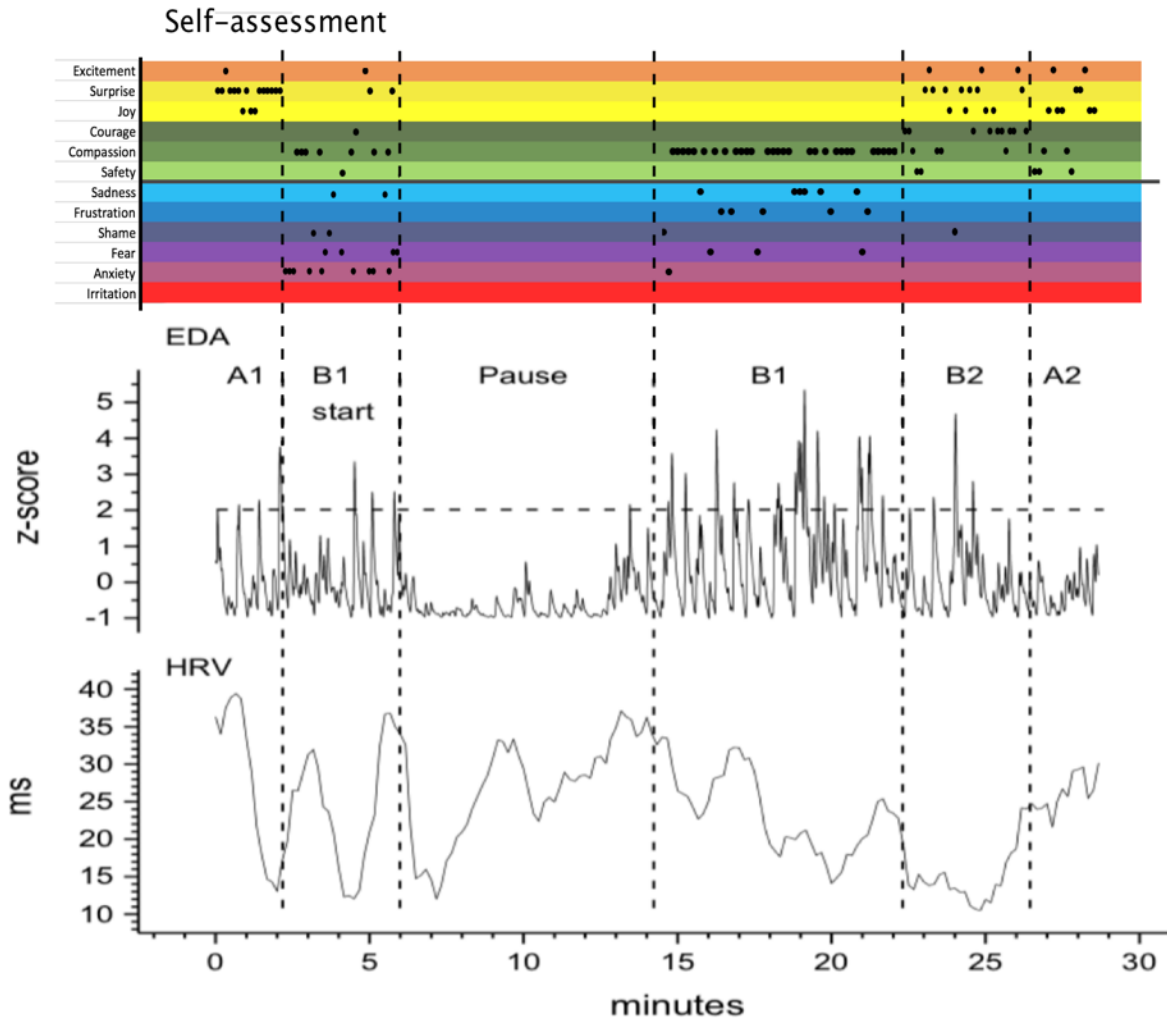

Figure 5. Self-assessments of emotions via EC, EDA, and HRV during Session I.

The decrease in the HRV continued further in episode B2 which was characterized by positive valence, but also by high EDA peaks. In the subjective self-assessment of episode B2, Lisa gave many reports of *surprise* together with *courage* (see Figure 5). However, towards the end of session B2, her HRV started to increase.

Here it should be noted that episode B2 was much shorter (4 min.) than episode B1 (9 min. 23 sec.). Since activation of parasympathetic nervous system combined with relaxing and calming down (indicated by the increase in HRV) takes place fairly slowly, this might have an influence on the delay in the HRV increase when HRV was activated within episode B2.

Next subchapter addresses to what extent does the emotion-driven Stimulated Recall Interview (SRI) promote reflection and hence learning, related to the original learning situation.

6.2. Findings from the Stimulated Recall Interview (SRI)

The assessments given via Emotion Circle (EC) together with the video-recorded learning episodes were used as the stimulus material for the SRI. In the SRI, we asked Lisa to explain her thoughts, feelings, and bodily sensations, plus her reasons for her assessments via EC. With this procedure (SRI) we mainly aimed to elicit the connections between emotions and the video-recorded learning situations. Our findings, based on a qualitative content analysis of the transcribed interview data, showed that the emotion-driven SRI produced very rich and heterogeneous descriptions, comprising self-reflection, other-related reflection, new emotion words, plus comments concerning the SRI method. These descriptions illustrated reasons for specific emotions, and the connections between the reported emotions and the details of the learning situations and processes. In addition, the SRI produced new insights concerning the meanings given by Lisa to her own past behaviour, as well as the change in Lisa's behaviour during the learning assignment. Table 2 provides a summary of the findings concerning the four episodes (A1, B1, B2, A2), including also a description of the emotional tone during the SRI.

A1 In the SRI for this episode, the emotions Lisa mentioned most frequently were 'surprise' and 'joy'. She commented on the reasons behind these emotions in a neutral and calm manner. She had felt surprised and happy at seeing many familiar individuals arriving at the training venue. There were no intense emotional expressions or reflective comments relating to this episode.

B1 In the SRI for this episode, Lisa provided nuanced descriptions, involving: (i) reasons for specific emotions, (ii) new emotions (i.e. emotions which she had not selected in the EC). She also gave reflective comments on the interactions while she viewed them. The comments involved (i) self-reflection on her own activities and bodily experiences, (ii) reflection on the activities of the other participants and the trainer, (iii) interpretative comments concerning other participants' behaviour, (iv) reflections on the group atmosphere.

In this episode, one participant was talking about her health problems. While watching the video, Lisa re-lived strong emotions, i.e. 'anxiety', 'sadness', and 'fear'. She described her re-lived emotions (and the self-reflection connected to these) as follows:

> …now I'm getting very anxious and I feel compassion, because *this is my group…*and *I'm responsible for the fact that they have worked too hard,* and I feel compassion and guilt, which is not included as an option in the EC. I swing between anxiety and compassion, but guilt is the strongest feeling, since *I truly see that people are working too hard, but I wonder how…I haven't realized that* [sighing] so that it feels really bad…that I can feel I've failed… and I can feel sadness…

Lisa also described embodied experiences, in comments such as 'surely, at that point my electrodermal activity was at a high level', 'I tried to take a deep breath in that situation', or 'there I definitely have tears in my eyes'.

Lisa suggested the addition of a new emotion word, 'guilt', which was not available in the EC. In addition, within the SRI, she named the emotions 'irritation', and 'joy', which she had not selected in the EC assessment phase, even though these emotion words had been options in the EC.

In the SRI for B1, Lisa also reflected on the behaviour of the group, as follows:

> ….then I look at the entire group, *the way everybody is sort of frozen,* or fortunately, it provides space [for the emotional expression] and *in this sense there is safety in the group, so that people can empathize with each other and feel relieved...* [smiling]

B2: This episode consisted of working with 'a difficult case' using a drama method. In the related SRI, Lisa provided diverse descriptions, including (i) reasons for the specific emotions, (ii) self-reflection and self-analysis, (iii) reflection on the group, and (iv) a new insight concerning the reasons why she herself perceived the case as so difficult.

The corresponding SRI started with Lisa's self-reflection and self-analysis connected to the selected EC emotion of 'surprise', and to how this surprise emerged from her role-playing of a difficult character. She described this as follows:

> …so I'm wondering and I'm surprised, wondering if that's me, I'm so bad at role-playing, but OK, I'm surprised that it's as if oh my god…*I have that difficult person in my mind all the time the one I'm role-playing…but also I'm astonished, about whether it's me that's speaking there or whether it's my role character*….actually I'm a bit ashamed about whether I'm so bad at role-playing but since I really am bad at role-playing…

In the citation above, there is also comment on her own behaviour and on issues which are bothering her. These troubling thoughts emerged especially in relation to the other group members, and what they might think about her own role-playing of a difficult person. This further created self-doubt concerning what she had said in the group. She expressed the troubled feelings, including reflection on the group, as follows:

> …and then I look at those group members *wondering what they're thinking because at that time* [referring to the original group learning situation] *I couldn't see it while I was concentrating on pretending to be that role character,* so what I was saying … I've really got such a that.. uh uh that that somehow I've spoken inappropriately…

After this, Lisa spoke in a way that implied balancing between the feeling of being troubled and the emotion of being courageous (the emotion she had chosen most in the EC). She vacillated between a feeling of having been courageous, and a feeling of being ashamed of her courage. This can be seen as an attempt on her part to seek different interpretations of herself. There seems to be a struggle between, on the one hand, finding her own possible space, within which she can allow herself to be courageous, and on the other hand, social embarrassment, to the point of shame. The reflections here seem to involve tacit (previously unverbalized or unrecognized) emotion. However, within the SRI, her talk took a positive direction, ending in laughter. In this way, the balance swung to a positive feeling, and to the conclusion that in fact she had not humiliated herself.

After this, a novel insight emerged on the reasons why Lisa had perceived the difficult case as so difficult. This insight emerged together with the reflective talk on her emotions, plus the reasons for these. In the SRI (while watching the situation again in the video and elaborating her emotion of being courageous), she pondered on her behaviour as follows:

> … I'm surprised at how courageous I am [in this learning assignment]. Um, I was pondering a lot …about whether I was going to dare to say it… that until then I was, well, that she was once my teacher…. when I started my studies in the Department… and now [in this learning assignment] I could at last dare to tell her … even though she had been an authority figure to me… well okay I was really courageous… I still keep wondering how and why I was so bold to as say to her that here we are all in the same problematic situation, so why, I guess she had put herself above all the others, and she was a university lecturer when I first started to work in the Department. *Well this has been for me I guess a kind of empowering moment and a really big issue, that I was able to tell her what I was thinking, since I have had a tendency to try to please everyone, and especially her, so that* [earlier, when we worked together] *I did not, I didn't dare* [laughs]…

Slightly later in the SRI, Lisa confirmed her new insight (which only emerged during the SRI) as follows:

> ...now I look at this from a distance, this is how I have been acting… well maybe *there's some new insight about why this person was for me, well since she had previously been a great authority for me then that's why it was such a big thing that after twenty or thirty years I had the courage to say to her you can't act like that,* tell people they can't come to a group if they haven't done their doctoral dissertation, just for that

> reason, come and make other people depressed, say you people are stupid [laughs] yeah this was something to remember…

Lisa also confirmed that this new idea, and the emotions attached to it, were merely those that emerged in the present situation (EC plus SRI). She could remember that in the original learning setting (five years previously) she had role-played that person, but she did not remember the emotions she experienced at that time.

A2: The SRI for this episode produced utterances that were descriptive and fairly neutral – but also moderately positive – concerning the group atmosphere. They encompassed feelings of being safe, joyful, and surprised. In addition, she presented descriptive comments concerning individual participants in the group.

Table 2
Summary of findings from the Stimulated Recall Interview (SRI) concerning the four episodes A1, B1, B2, and A2

Episode	New emotions mentioned in the SRI (not selected in the EC)	Emotional tone during the SRI	Self-reflective comments in the SRI	Other-related reflection in the SRI	New learning and insight in the SRI	Comments on the SRI as an experience
A1		Neutral valence Reasons for selected emotions given, no intense emotional expression during the SRI				
B1	Not available in the EC: • guilt Available in the EC, not selected: • frustration • irritation • joy	Negative valence Emotional expression • sadness (tears in her eyes) Emotions re-lived • anxiety • fear • sadness • guilt	Self-reflection on the subject's • role as a leader Critical reflection on feeling herself as personally responsible for not being observant enough	Reflection on: • other group members • trainer • entire group		Emotion mentioned: • relief SRI was a therapeutic experience for the subject
B2		First mixed, then positive valence Emotions expressed: • joy (laughter)	Self-reflection on the subject's • behaviour while adopting the role of "a difficult person" • skills in presenting the role • being brave in the assignment Self-analysis: presenting herself as brave vs. being ashamed of it	Reflection on: • entire group	New learning and insight: The subject realized the reasons behind the problematic relationship addressed in the episode	The subject confirmed the insight she had gained in the SRI
A2		Positive valence Emotions expressed: • joy				Emotionally, A1 and A2 were not as intense as B1 and B2 for the subject

7. Conclusions regarding complementarity and further challenges in using multiple methods

Our aim in this pilot study was to gain a preliminary understanding of how self-reports, indicators of ANS, behavioural measures of emotion, and the emotion-driven Stimulated Recall Interview (SRI) may provide complementary information on the function of emotions in professional learning. In conducting the study, we wished also to elaborate the potentials and challenges of the multi-componential methodology we applied, in terms of researching emotions in professional learning.

Our interest in investigating emotions in relation to learning processes implies a need for a bi-directional perspective, involving how emotions elicit learning processes and outcomes, and how learning processes elicit emotions. In future discussion on emotions in learning we also need to consider recent discussion and disagreement between, on the one hand, scholars who think that emotions are universal and similar over different times and cultures (Ekman, 1992), and on the other hand, those who see emotions as historically and culturally determined, and thus learned in socio-cultural contexts (Barrett, 2006). The methods and tools used to measure emotions from facial expressions (such as the FaceReader used in this study) based on Ekman's (1992) idea of basic emotions and the possibilities for universal measurement of them via facial movements. However, this may be misleading, bearing in mind the criticisms presented against the conception of universality in facial movements as indicators of emotions. Human capabilities for emotion regulation, and individual differences in emotional intelligence (manifested as the ability to display emotions), can be expected to influence the presentation of emotions. There is, in fact, considerable evidence concerning the learning of emotions in cultural contexts, and this applies also to the learning and use of emotion words (Barrett, 2006).

The on-line self-reports (given via EC) and the SRIs based on these emerged as productive, not only in reporting and explaining one's emotions during the learning process, but also in terms of promoting new reflective learning. This new learning was evidenced in what our subject, Lisa, said in her SRIs, and especially in her elaborations of the emotions experienced while viewing the episodes.

Lisa's self-reports via the EC, and behavioural data derived from her facial expressions, indicated that the selected episodes B1 and B2 were both connected with intense emotions. They also indicated a difference in valence, with the first of these (B1) being characterized by negative valence, and the second (B2) by positive valence. This was evident on the basis of her self-reports, and was validated by behavioural data derived from her facial expressions. The findings here include considerable measurement error, due to our subject's need to use glasses. Nevertheless, they are in line with previous studies showing high agreement between facial expressions and self-report data (Harley et al., 2015).

In the SRI, Lisa elaborated on her reasons for her emotion choices in the EC. While viewing the video clips, she could explain what had provoked the emotions. In the episode imbued with emotions of negative valence (B1), Lisa reported compassion, anxiety, sadness, and fear. The next episode (B2), which was characterized by emotions with positive valence (surprise, courage, joy, compassion, excitement, and safety) seemed to produce new self-reflective ideas on the previous power relations operating between herself and a particular difficult person, and on why the relationship had been so difficult for her. This finding is in line with previous suggestions regarding the SRI method as a means of promoting self-reflection (Vall et al., 2018). However, the special feature in the present study was the use of the SRI to focus specifically on emotions. The emotion-driven SRI seemed to produce – in conjunction with reflections by Lisa on herself, other members, the group interaction, and the reasons for her specific emotions – important insights concerning the causes underlying her 'difficult case'.

Such novel learning, involving new insights, seems to support the productive nature of *emotion-driven reflection on personally meaningful learning episodes*. While strong emotions may exist as (partly unconscious) rapid events at particular moments, the elaboration of these moments afterwards, i.e. from a distance, appears to provide a productive basis for identity learning.

The video-recorded episodes from personally meaningful learning settings unfolded first from distance in watching and assessing these via EC, and then slowly in the SRI. They thus provided options to re-live and re-analyse the connections between the situation in which the emotions arose and the emotional responses that followed (see Figure 1). All in all, this implies that within a methodology of presenting stimulus material, it is

highly productive to select personally meaningful episodes for further elaboration, and to elaborate these from the perspective of emotions.

The self-report method used in this study (combining the on-line EC reporting of emotions with the SRI focusing on personally meaningful learning episodes) provided a powerful learning setting. It brought about deepened self-reflection, and novel insights, notably in the episode characterized by emotions with positive valence. Overall, we would suggest that a combination of the on-line reporting of emotions with the emotion-driven SRI is a promising methodology for investigating how emotions are connected to learning. In addition, such a combination can promote reflective insights that are of value in learning about one's own identity. There is thus notable pedagogical potential in the SRI, which appeared to bring about self-reflection, other-oriented reflection, insight, and learning. This finding resonates with the studies by Vall et al. (2018) and Huhtamäki et al. (2017), who noticed that the video-assisted SRI stimulated reflection and insight, while facilitating therapeutic processing in couple therapy clients.

During the learning process in session 1, the psychophysiological indicators EDA and HRV were found to provide different and complementary information on the subject's autonomic nervous system (ANS) activity. The EDA, which indicates the activity level of the sympathetic nervous system, was found to be at a high level during the viewing and assessing of the videotaped learning episodes. The lowest EDA occurred during the pause, i.e. during a passive situation when there was nothing to do. Comparison of the EDA responses between the watched episodes showed that the highest EDA peaks occurred during the viewing of personally significant episodes B1 and B2. In episode B1 (characterized by emotions with negative valence), there were many significantly high peaks, and the peaks were close to each other (i.e. dense). Nevertheless, high EDA peaks were also present in the other personally significant episode, i.e. the one characterized by positive emotions (B2). In this study, EDA could be used to distinguish the personally meaningful episodes B1 and B2 from the passive and more neutral episodes A1 and A2. However, EDA did not distinguish between the emotional valences (positive vs. negative) of the situation. For this purpose, we need other complementary methods, such as self-reports.

Another ANS measure used here was heart rate variability (HRV), which is a measure of the activity of the parasympathetic nervous system (PNS). In the present study, HRV increased (indicating high activity in the PNS) thus pointing to processes of calming and relaxing during the (unintended) pause, as well as during emotionally neutral episodes A1 and A2. In contrast, HRV was found to decrease (thus indicating deactivation of the PNS) in the emotionally intense episodes B1 and B2. One very interesting aspect was the decrease in HRV also during the transition from episode B1 (imbued with negative emotions) to episode B2 (imbued with positive emotions). However, the findings here should be treated with caution, given that this is a single-subject case study, and also that sensitivity to breathing (which is a feature of HRV) could have affected the HRV observations. In the present study we were unable to obtain breathing data due to a technical failure. Hence, for future collection and analysis of HRV data it will be necessary to improve the reliability of the relevant equipment.

The behavioural data (via face reader, gaze) were used merely to increase the reliability of the self-reporting data. In the face expression analysis we utilized – in line with our dimensional understanding of emotions – only the indicator of valence. As regards gaze data, in the present study these data were collected but not further analysed. However, the collection of such data would have utility in showing the focus of attention. In the future, we intend to use gaze data in addressing further the usability of EC.

Our efforts to construct measurement procedures for use under laboratory conditions produced many insights regarding technical details. These covered the proper installation of devices (such as, in the present case, the breathing belt which did not give proper data because it was too loosely tightened). Furthermore, from the unplanned loss of the sound in the video display, and from the ANS data collected during the pause, we learned that we actually *require such a pause situation* at the end of data collection sessions, to obtain the baseline for the subject's ANS data.

One central challenge in this kind of multimethod measurement is the synchronisation of different devices. If there are problems with this, analysis of the data becomes very laborious. Furthermore, in analysing ANS data there are different time windows in EDA and HRV. EDA, which measures the responses of rapid

sympathetic response systems, responds quickly (e.g. in a fight-flight situation). This means that the EDA response occurs in just a few seconds from the stimulus situation or event. By contrast, the HRV measures the activity of the (slowly responding) parasympathetic nervous system. This means that in counting the HRV, the time window cannot be less than one minute. Hence, the temporal connection to a specific event or situation remains less exact with HRV as compared to EDA.

For the future development of the Emotion Circle (EC), we gained much information from the emotion-driven Stimulated Recall Interview (SRI). It provided practical information concerning the usability and selection of emotion words. For the future, we need to test the optimal number of emotion words, and also how to take into account the saturation of colours in EC. In addition, there is a need to test a range of pictorial or iconic ways of displaying emotions in the EC.

In the future, if our purpose is to measure simultaneously all three components of emotions (subjective experience, ANS, and behaviour) within the processes of learning, we shall need a multidisciplinary team. We thus need to recognize that setting up this kind of multi-method measuring system will require a multidisciplinary group of researchers, comprising experts in educational sciences, psychology, psycho-physiology, and information technology. In addition, there will be a need for different kinds of inter-professional practical support and technical services.

For future research on emotions in learning, we would suggest a focus on the continuities of emotional processes (in terms of the bodily activity of the ANS, plus valence, and the intensity of the experienced emotions). The aim will be to create conditions that are optimal for researching and promoting supportive emotions in professional learning.

Keypoints

- This study developed a multi-componential methodology to measure emotions in learning.
- An on–line assessment tool, Emotion Circle (EC), was developed for the self-reporting of emotions during learning.
- The multimethod research design provided complementary information on the experiential, physiological, and behavioural components of emotions.
- The emotion-driven SRI revealed connections between emotions and learning, and was productive in bringing about reflective learning and novel insights.

Acknowledgments

The authors are grateful to the reviewers of the manuscript, and to Donald Adamson who polished the language of the article. They also wish to thank Lauri Viljanto and Petri Kinnunen for technical assistance, and Hanna Lilja who drew the figures. This work was supported by the Academy of Finland [Grant number 288925 The Role of Emotions in Agentic Learning at Work].

References

Azevedo, R., Harley, J., Trevors, G., Duffy, M., Feyzi-Behnagh, M., & Bouchet, F. et al. (2013). Using trace data to examine the complex roles of cognitive, metacognitive, and emotional self-regulatory processes during learning with multi-agent systems. In R. Azevedo & V. Aleven (Eds.), *International Handbook of Metacognition and Learning Technologies* (pp. 427-449). New York: Springer.

Azevedo, R., Taub, M., Mudrick, N. Farnsworth, J. & Martin, S. (2016). Interdisciplinary research methods used to investigate emotions with advanced learning technologies. In M. Zembylas & P. Schutz (Eds.), *Methodological Advances in Research on Emotion and Education* (pp. 231-243). Dordrecht: Springer.

Barsade, D., & Gibson, S. G. (2007). Why does affect matter in organizations? *Academy of Management Perspectives*, *21*, 36–59.

Barrett, L. F. (2006). Are emotions natural kinds? *Perspectives on Psychological Science*, *1*, 28-58.

Barrett, L. F. & Russell, J. A. (1999). The structure of current affect. Controversies and emerging consensus. *Current Directions in Psychological Science, 8,* 10–14.

BeGaze Version 3.7. (2016). *Manual.* SMI SensoMotoric Instruments.

Benedeck, M. & Kaernback, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*, *190*, 80-91.

Bondareva, D., Conati, C., & Feyzi-Behnagh, R. (2013). Inferring learning from gaze data during interaction with an environment to support self-regulated learning. *International Conference on Artificial Intelligence in Education* (pp. 229-239). Springer.

Boucsein, W. (2012). *Electrodermal activity.* New York: Springer.

Brentson, G. G., Bigger, Jr., Eckberg, D. L., Grossman, P., Kaufman, P. G., Malik, M., Nagaraja, N. H., Porges, S. W., Saul, J. P., Stone, P. H. & Van Den Molen, M. W. (1997). Heart rate variability: Origins, methods, and interpretive caveats. *Psychophysiology*, *34*, 623-648.

Darwin, C.R. (1872/1965). *The expression of emotions in man and animals*. London, UK: John Murray.

Dimaggio, G., Lysaker, P., Carcione, A., Nicolo, G., & Semerari, A. (2008). Know yourself and you shall know the other to a certain extent: Multiple paths of influence of self-reflection on mindreading. *Consciousness and Cognition*, *17*, 778-789.

Ekman, P. (1992). Argument for basic emotions. *Cognition and Emotion*, *6*(3-4), 169-200.

Eteläpelto, A. (2017). Emerging conceptualisations on professional agency and learning. In M. Goller, & S. Paloniemi (Eds.), *Agency at work: An agentic perspective on professional learning and development* (pp. 183-201). Springer: Cham.

Eteläpelto, A., Hökkä, P., Paloniemi, S., Vähäsantanen, K., Lappalainen, V., Eteläpelto, T., & Niskanen, K. (2017). Developing an on-line application, Emotion Circle (EC) for the self-assessment of emotions in agentic learning at work. In L. G. Chova, A. L. Martínez, & I. C. Torres (Eds.), *ICERI 2017 Proceedings:10th International Conference of Education, Research and Innovation.* Seville, Spain (pp. 7763-7769).

Eteläpelto, A., Vähäsantanen, K., Hökkä, P. & Paloniemi, S. (2013). What is agency? Conceptualizing professional agency at work. *Educational Research Review, 10,* 45-65. http://authors.elsevier.com/sd/article/S1747938X13000274

Eteläpelto, A., Vähäsantanen, K., Hökkä, P. & Paloniemi, S. (2014). Identity and agency in professional learning. In S. Billett, C. Harteis and H. Gruber (Eds.), *International Handbook of Research in Professional and Practice-based Learning Volume 2* (pp. 645-672). Dordrecht: Springer.

FaceReader Version 6.1. (2015). *Reference Manual.* Wageningen, NL: Noldus Information Technology.

Fredrickson, B. L., & Branigan, C. (2002). Positive emotions broaden the scope of attention and thought-action repertoires. *Cognition and Emotion*, *16*, 313-332.

Frijda, N. H. (1986). *The Emotions*. Cambridge, UK: Cambridge University Press.

Fuller, F. F., & Manning, B. A. (1973). Self-confrontation reviewed: A video tape. A case study. *Journal of Counseling Psychology*, *10*(3), 237-243.

Gendolla, G. H. E. (2017). Do emotions influence action? Of course, they are hypo-phenomena of motivation. *Emotion Review*, 9(4), 348-350.

Harley, J.M., Bouchet, F., Hussain, M.S., Azevedo, R. & Calvo, R. (2015). A multi-componential analysis of emotions during complex learning with an intelligent multi-agent system. *Computers in Human Behavior, 48,* 615-625.

Hautala, J., Loberg, O., Hietanen, J. K., Nummenmaa, L., & Astikainen, P. (2016). Effects of conversation content on viewing dyadic conversations. *Journal of Eye Movement Research, 9*(7), 1-12.

Hökkä, P., Vähäsantanen, K., Paloniemi, S., & Eteläpelto, A. (2017). The reciprocal relationship between emotions and agency in the workplace. In M. Goller, & S. Paloniemi (Eds.), *Agency at work: An agentic perspective on professional learning and development* (pp. 161-181). Springer: Cham. doi:10.1007/978-3-319-60943-0_9

Homan, A.C., Van Kleef, G.A. & Sanchez-Burks, J. (2015). Team members' emotional displays as indicators of team functioning. *Cognition and Emotion, 30*(1), 134-149.

Hommel, B., Moors, A., Sander, D., & Deonna, J. (2017). Emotion meets action: Towards an integration of research and theory. *Emotion Review*, *9*(4), 295-298.

Huhtamäki, H., Lehtinen, R., Kykyri, V-L., Penttonen, M., Karvonen, A., Kaartinen, J., & Seikkula, J. (2017). Oivaltamisen hetket pariterapia-asiakkaiden jälkihaastatteluissa.[Moments of insight in the stimulated recall interview of couple therapy clients]. *Perheterapia*, *33*(4), 36-50.

Hugdahl, K. (1995). *Psychophysiology: The mind-body perspective.* Cambridge, MA: Harvard University Press.

James, W. (1884). What is emotion? *Mind, 9,* 188-205.

Kagan, N., Krathwohl, D. R., & Miller, R. (1963). Stimulated recall in therapy using video tape: A case study. *Journal of Counseling Psychology, 10,* 237-243.

Karvonen, A. (2017). Sympathetic nervous system synchrony between participants of couple therapy. *Jyväskylä Studies in Education, Psychology and Social Research* 599. University of Jyväskylä. Jyväskylä University Printing House, Jyväskylä, Finland.

Kratochwill, T.R., Hitchcock, J.H., Horner, R.H., Levin, J.R., Odom, S.L., Rindskopf, D.M., et al. (2013). Single-case intervention research design standards. *Remedial and Special Education, 34,* 26-38.

Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: A review. *Biological Psychology, 84*(3), 394–421.

Kykyri, V.-L., Karvonen, A., Wahlström, J., Kaartinen, J., Penttonen, M., & Seikkula, J. (2017). Soft prosody and embodied attunement in therapeutic interaction: A multimethod case study of a moment of change. *Journal of Constructivist Psychology, 30* (3), 211-234. doi:10.1080/10720537.2016.1183538

Labone, E., & Long, J. (2016). Features of effective professional learning: A case study of the implementation of a system-based professional learning model. *Professional Development in Education, 42*(1), 54–77.

Laitila, A., Vall, B., Penttonen, M., Karvonen, A., Kykyri, V-L., Tsatsishvili, V., Kaartinen, J. & Seikkula, J. (2018). The added value of studying embodied responses in couple therapy research: A case study. *Family Process 57,* 1-13.

Lazarus, R. S. (1991). *Emotion and adaptation*. New York: Oxford University Press.

Levenson, R. W. (1994). Human emotion: A functional view. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 123-126). New York: Oxford University Press.

Levenson, R. W. (2003). Autonomic specifity and emotion. In R. J. Davidson, K. R. Scherer & H. H. Goldschmith (Eds.), *Handbook of affective sciences* (pp. 212-224). New York: Oxford University Press.

Levenson, R. W. (2014). Autonomic nervous system and emotion. *Emotion Review*, *6*(2), 100-112.

Lyle, J. (2003). Stimulated recall: A report on its use in naturalistic research. *British Educational Research Journal*, 29(6), 861-878.

Manolov, R. & Onghena, P. (2017). Analyzing data from single-case alternating treatment designs. *Psychological Methods*, *16*(March), 1-25.

Mauss, I. B., Levenson, R. W., McCarter, L., Wilhelm, F. H., & Gross, J. J. (2005). The tie that binds? Coherence among emotion experience, behaviour and physiology. *Emotion*, *5*, 175-190.

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition and Emotion, 23,* 209-237.

Pekrun, R. (2016). Using self-report to assess emotions in education. In M. Zembylas, M. & P. A. Schultz (Eds.), *Methodological advances in research on emotion and education.* Springer International Publishing.

Pekrun, R., Elliot, A. J., & Maier, M. A. (2006). Achievement goals and discrete achievement emotions: A theoretical model and prospective test. *Journal of Educational Psychology*, *98*(3), 583-597.

Perry, D.P. (2006). Fear and learning. Trauma-related factors in the adult education process. *New Directions for Adult and Continuing Education*, *110*, 21-27.

Philpott, C., & Oates, C. (2017). Teacher agency and professional learning communities: What can learning rounds in Scotland teach us? *Professional Development in Education*, *43*(3), 318–333.

Porges, S.W. (2001). The polyvagal theory: Phylogenetic substrates of a social nervous system. *International Journal of Psychophysiology*, *42*, 123-146.

Posner, J., Russell, J. A. & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, *17*(3), 715-734.

Quintana, D. S. & Heathers, J. A. J. (2014). Considerations in the assessment of heart rate variability in biobehavioral research. *Frontiers in Psychology*, *5*, 805. doi:10.3389/fpsyg.2014.00805

Quintana, D. S., Alvares, G. A. & Heathers, J. A. J. (2016). Guidelines for reporting articles on psychiatry and heart rate variability (GRAPH): recommendations to advance research communication. *Transl. Psychiatry*, *6*, 803.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*, 1161–1178.

Russell, J. A. (1989). Affect grid: A single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology*, *57*(3), 493-502.

Russell, J. A. (2003). Core affect and the psychological construct of emotion. *Psychological Review*, *110*(1), 145-172.

Russell, J. A. (2005). Emotion in human consciousness is built on core affect. *Journal of Consciousness Studies, 12*(8-10), 26-42.

Russell, J. A., Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology, 76*(5), 805-819.

Scherer, K. R. (2009). The dynamic architecture of emotion: Evidence for the component process model. *Cognition and Emotion, 23,* 1307-1351.

Stein, P. K., Bosner, M. S., Kleiger, R. E., & Conger, B. M. (1994). Heart rate variability: A measure of cardiac autonomic tone. *American Heart Journal, 127,* 1376–1381.

Storbeck, J. & Maswood, R. (2015). Happiness increases verbal and spatial working memory capacity where sadness does not: Emotion, working memory and executive control. *Cognition and Emotion*, *30,* 925–938.

Sung, B., & Yih, J. (2015). Does interest broaden or narrow attentional scope? *Cognition and Emotion*, *30*(8), 1485-1494.

Tarvainen, M., Lipponen, J., Niskanen, J.-P., & Ranta-aho, P.O. (2018). *Kubios HRV (ver. 3.1.) User's guide*. Kubios Oy (limited company) /www.kubios.com/

Tomkins, S. S. (1962). *Affect, imagery, consciousness. Vol. 1. The positive affects*. New York: Springer.

Vähäsantanen, K., Hökkä, P., Paloniemi, S., Herranen, S., & Eteläpelto, A. (2017). Professional learning and agency in an identity coaching programme. *Professional Development in Education*, *43*(4), 514-536. http://www.tandfonline.com/eprint/FYuMGIBSbIIGezx5mxZP/full

Vähäsantanen, K., Paloniemi, S., Hökkä, P., & Eteläpelto, A. (2017a). Agentic perspective on fostering work-related learning. *Studies in Continuing Education, 39*(3), 251–267. http://www.tandfonline.com/doi/full/10.1080/0158037X.2017.1310097

Vähäsantanen, K., Paloniemi, S., Hökkä, P., & Eteläpelto, A. (2017b)**.** An agency-promoting learning arena for developing shared work practices. In M. Goller, & S. Paloniemi (Eds.), *Agency at work: An agentic perspective on professional learning and development* (pp. 351–371). Springer: Cham. doi:10.1007/978-3-319-60943-0_18

Vähäsantanen, K., Räikkönen, E., Paloniemi, S. Hökkä, P., & Eteläpelto, A. (2018). A novel instrument to measure the multidimensional structure of professional agency. *Vocations and Learning*, https://doi.org/10.1007/s12186-018-9210-6

Vall, B., Laitila, A., Borcsa, M., Kykyri, V.L., Karvonen, A., Kaartinen, J., Penttonen, M., & Seikkula, J. (2018). Stimulated Recall Interviews: How can the research interview contribute to new therapeutic practices? *Revista Argentina de Clinical Psicologica, XXVII, 2, 274-293,* DOI: 10.24205/03276716.2018.1068

Winkelman, P., & Berridge, K. C. (2004). Unconscious emotion. *Current Directions in Psychological Science*, *13*, 120-123.

Winkler, I. (2018). Identity work and emotions: A review. *International Journal of Management Reviews, 20,* 120–133.

Öhman, A., Carlsson, K., Lundqvist, D., & Ingvar, M. (2007). On the unconscious subcortical origin of human fear. *Physiology and Behavior*, *92*, 180-185.

Zembylas, M., & Schultz, P. A. (2016). (Eds*.), Methodological advances in research on emotion and education.* Springer International Publishing.

Zwart, R.C., Korthagen, F.A.J, & Attema-Noordewiever, S., (2015). A strength-based approach to teacher development. *Professional Development in Education*, *41*(3), 579–566.