

Article

Progression of Cognitive-Affective States During Learning in Kindergarteners: Bringing Together Physiological, Observational and Performance Data

Priyashri Kamlesh Sridhar * and Suranga Nanayakkara * 🝺

Augmented Human Lab, Auckland Bioengineering Institute, The University of Auckland, 1010 Auckland, New Zealand

* Correspondence: priya@ahlab.org (P.K.S.); suranga@ahlab.org (S.N.)

Received: 6 June 2020; Accepted: 25 June 2020; Published: 3 July 2020



Abstract: It has been shown that combining data from multiple sources, such as observations, self-reports, and performance with physiological markers offers better insights into cognitive-affective states during the learning process. Through a study with 12 kindergarteners, we explore the role of utilizing insights from multiple data sources, as a potential arsenal to supplement and complement existing assessments methods in understanding cognitive-affective states across two main pedagogical approaches—constructionist and instructionist—as children explored learning a chosen Science, Technology, Engineering and Mathematics (STEM) concept. We present the trends that emerged across pedagogies from different data sources and illustrate the potential value of additional data channels through case illustrations. We also offer several recommendations for such studies, particularly when collecting physiological data, and summarize key challenges that provide potential avenues for future work.

Keywords: kindergartners; cognitive-affective state; learning; empirical study

1. Introduction

As parents, educators and researchers, much of our understanding of and approach to how children learn is driven by our own beliefs and definitions of what it means to learn. As learners interact with the environment and learning content, there is a progression of cognitive-affective states as a response to learning. Such states also affect and determine learning and are influenced by the pedagogy [1]. While our beliefs drive our pedagogical approach, our assessments of learning are influenced by the pedagogies themselves. Irrespective of the approach towards learning, one commonality they share has been the complex interaction of cognitive and affective states during learning. Yet, the assessments of learning tend to focus majorly on achievement of curricular goals and skills that, while undoubtedly essential [2], present only one facet of something as complex as learning. Such assessments in education and learning are broadly categorized as diagnostic (assessment before teaching), formative (assessment during learning) and summative (assessment after teaching).

Summative assessments are conducted at the end of the lessons, making them too late for just-in-time rectification and remediation [3]. Formative assessments that include observations, conversations and discussions, questioning, peer assessments, self-reflection, essays, tests etc., demand huge amounts of time and planning that may sometimes affect the lesson time. Furthermore, while observations do offer some insights into cognitive-affective states of the learner, such as what concepts may be challenging, easy or boring, they are prone to biases and may not truly reflect states in learning [4,5]. It has been shown that combining a variety of data, including observations, responses, self-reports and physiological information [6], results in more powerful insights [7]. In fact,



with the advent of affective computing, physiological sensors are being used in the context of automated affect-aware tutors to infer affect and intervene accordingly [8]. However, most of the work has been conducted with older students and adults.

Drawing inspiration from earlier work [6], we use a combination of physiological, observational and performance measures to understand the patterns and progression of cognitive-affective states in kindergartners as they learn Science, Technology, Engineering and Mathematics (STEM) concepts through constructionist and instructionist approaches and explore key elementary differences across these approaches. Constructionism refers to educational practices that are student-focused, meaning based, process oriented, interactive, and responsive to student interest [9]; it believes that knowledge construction occurs when learners are engaged in building objects. Instructionism, in current educational contexts, refers to teacher-centered, highly structured, and non-interactive instructional practices [9]. It is also described as systematic teaching, explicit teaching, direct teaching, and active teaching [10] with a higher emphasis on the teacher than the student [11]. There has been a lot of impetus to implement STEM learning in kindergarteners between 5–8 years of age, as young children and scientists are often known to share a sense of innate curiosity that makes them want to know more about the world through experiment and observation [12]. Further, it has been shown that active engagement by K–12 students in hands-on science activities that use authentic science tools promotes student learning and retention of various skills, such as problem-solving, comprehension and communication skills, to name a few [13].

We use a mix of quantitative and qualitative approaches to appreciate the complexity of real-life learning experiences and share some of the insights and challenges in conducting such studies. Further, we do so in both constructionist and instructionist approaches that are two of the popular methods for learning and teaching, with a view to understand any differences in cognitive-affective state progression across the two approaches. The specific contributions include:

- Understanding and illustrating the progression of cognitive-affective states as kindergartners learn STEM concepts taught using constructionist and instructionist approaches through the collection and analysis of physiological, observational and performance measures.
- A synthesis of several recommendations based on our study and previous work for conducting similar studies with young children, especially the collection and analysis of physiological data.

1.1. Background

1.1.1. Cognitive Affective States in Learning

Learning involves a complex coordination of cognitive processes and affective states. Cognitive processes that underlie inference generation, causal reasoning, problem diagnosis, conceptual comparisons, and coherent explanation generation are accompanied by affective states such as irritation, frustration, anger and sometimes rage when the learner makes mistakes, struggles with troublesome impasses and experiences failure [5]. On the other hand, positive affective states such as flow, delight, excitement and eureka are experienced when tasks are completed, challenges are conquered, insights are unveiled and major discoveries are made. Simply put, emotions are systematically affected by the knowledge and goals of the learner, as well as vice versa [14]. The inextricable link between affect and cognition is a fundamental assumption adopted by the major theories of emotion.

Some researchers have integrated both affective and cognitive components into motivation theories [15–17]. While these have provided insight into emotions and reaffirmed their role in learning, there is not much consensus on the kind of emotions involved in learning. Csikszentmihalyi [18] has emphasized the tendency for a pleasurable state of "flow" to accompany problem solving. Kort et al. [19] attempted to list the emotions involved in learning and proposed a four quadrant model relating phases of learning to emotions. Pekrun et al. [20] classified academic emotions into three categories including achievement emotions, topic emotions and epistemic emotions.

Research has focused on a more in-depth analysis of a smaller set of emotions that arise during deep learning in more restricted contexts and over shorter time spans, from 30 to 90 min [5,8,21,22].

The prominent emotions in these learning sessions include boredom, engagement, confusion, frustration, anxiety, curiosity, delight, and surprise [23]. While it has been acknowledged that the identification of the affective states during learning is critical, merely knowing the state has limited use. As D'Mello et al. [5] state, what is missing is a specification of how these states evolve, morph, interact and influence learning and engagement, and this is particularly limited for younger age groups in learning situations.

1.1.2. Assessments in Learning

There has always been a growing recognition amongst teachers and educators that student interest, motivation and participation are highly important in learning [24,25]. While these correspond with students' cognitive-affective states, much of these judgements come from intuition and therefore do not necessarily feature as the priority in school assessments. This section addresses some of the most commonly employed assessment tools, what they are typically used to measure and their advantages and drawbacks.

Assessment is the process of gathering information that accurately reflects how well a student is achieving the curriculum expectations in a subject or course [3]. The primary purpose of assessment is to improve students' learning. "It is worth noting, right from the start, that assessment is a human process, conducted by and with human beings, and subject inevitably to the frailties of human judgement. However crisp and objective we might try to make it, and however neatly quantifiable may be our results, assessment is closer to an art than a science. It is, after all, an exercise in human communication" [26] (p. 2).

Assessments in the current educational scene are broadly categorized as diagnostic, formative and summative assessments. However, as Harlen [27] explains: "Using the terms 'formative assessment' and 'summative assessment' can give the impression that these are different kinds of assessment or are linked to different methods of gathering evidence. This is not the case; what matters is how the information is used. It is for this reason that the terms 'assessment for learning' and 'assessment of learning' are sometimes preferred. The essential distinction is that assessment for learning is used in making decisions that affect teaching and learning in the short term future, whereas assessment of learning is used to record and report what has been learned in the past" (p. 104). Therefore assessments can be categorized based on the nature of use of the information collected. Accordingly, diagnostic assessment would be used as assessment for learning before instruction, formative assessment can be used as both assessment for learning and assessment as learning during instruction, and summative is used as assessment of learning after instruction where the goal is evaluation.

- 1. *Diagnostic/assessment for learning*: The teacher gathers information on the student's existing skills, preferences, learning needs and readiness. The information is then used to plan the lessons and learning paths that are differentiated and suited to the learner's needs and skills.
- 2. Formative/assessment for learning: The teacher conducts the assessments to evaluate learners' skills and the nature of students' strengths and weaknesses, and to help develop students' capabilities. The information is collected in a planned manner during ongoing learning and used to keep track of the progress with respect to the curricular goals and further set individual goals for learning. These may include formal and informal observations, discussions, learning conversations, questioning, conferences, homework, tasks done in groups, demonstrations, projects, portfolios, performances, peer and self-assessments, self-reflections, essays and tests [3].
- Summative/assessment of learning: These assessments occur at the end of a learning period and are designed to provide information about student achievement and also work as an accountability mechanism to evaluate teachers and schools [28]. These include performance tasks, exams, standardized exams, demonstrations, projects and essays.

Some of the commonly used assessment tools, such as tests and examinations, self-reports, teachers' ratings and observations, are used to infer cognitive load imposed by learning or the mental effort exerted by the learner [29,30]. While these assessment are valuable and provide different

dimensions of learning and the learner, relying on just one of them or on the tools that are currently in practice brings with it certain challenges. In general, all assessments require a great deal of thought and planning to be effective. One of the most commonly employed assessment tools consists of tests and examinations. While they are time and cost effective and have limited scope for unfairness and biases, they are conducted after learning and do not really influence much of subsequent teaching for the same child. Performance measures such as tests and exams, projects etc., are ridden with some fundamental challenges. Tests encourage rote and superficial learning and questions are not critically renewed with respect to what they assess. There is also a tendency to emphasize quantity and presentation of work and to neglect its quality in relation to learning. There is an over-emphasis on the cognitive aspects of learning (memory and retrieval) instead of a more balanced emphasis on cognitive and affective aspects of learning despite prevalent support for both these states being impactful on learning.

Another commonly employed assessment tool is observation. Experienced teachers rely a lot on observations of their students to understand what students may be experiencing, their performance, their emotions and their affective states, and use their observations as part of their formative and summative assessment [31]. These observations can be incidental (teacher observes the student during the teaching and learning and may use this as part of the formative assessment) or planned (the observation is planned by the teacher to observe specific learning outcomes and this may be part of a formal assessment task). Observations provide a great deal of information about student learning outcomes, especially in early childhood education. Most importantly, they offer a more continuous access to a student's affective states, as teachers can interpret affective behaviors, such as emotional state, posture and gaze, by observing the learner in an appropriate context. However, observations are wrought with challenges too. Firstly, the point at which the observations are carried out may not offer the student an opportunity to demonstrate a skill. Secondly, observing a class of students consistently is challenging given the sheer number of students, the number of tasks to be accomplished by the end of the class/term and other behaviors from specific students that need more attention. Thirdly, observations are influenced by prior judgements and prejudices. Finally, certain states, such as boredom, confusion or anxiety, may not always have overt manifestations, which means the teacher may erroneously misinterpret a child's behavior or miss one.

The third most common form of assessment include self-reports. These are powerful in that they offer the learner a chance to reflect on their learning and share their opinions and experiences, which ideally helps teachers to guide their teaching through an understanding of their pupils. A lot of the survey and self-report methods rely on the question–answer process. While this may be oral or written, asking good questions is challenging, and further, understanding and interpreting that question and formulating an appropriate response can be even harder for some children. Such questions can be open- or close-ended. While close-ended questions are slightly easier, open-ended ones are harder to interpret. Further, self-reports for children below age 11 have low validity due to their limited language ability, reading age, motor skills and temperamental effects, such as confidence, self-belief and desire to please [32]. Most of the studies present the questionnaires after the learning has occurred [33,34]. As a result, there is a high possibility that the participant may provide an average estimate for the whole task that is affected by memory effects. This loses its purpose of capturing the dynamic and fluctuating nature of the load that is imposed during learning [35].

With the inclusion of technology into learning environments, there have been some encouraging explorations on human thinking and information processing abilities [25]. Many of the studies of cognitive load and learning outcome measures administer each of the above measures either before or after test performance [36]. Even though they are static and considered unreliable [37], they continue to be popular in real-world contexts partly because single measures are easy to administer whereas other objective measures of cognitive load may require expensive and hard-to-use instrumentation. By contrast, Mayer et al. [38] have highlighted the need for direct measure of cognitive overload.

Physiological measures such as Electro-Dermal Activity (EDA) and Heart-Rate Variability (HRV) offer a direct, sensitive and usable measure of cognitive load [25,39,40].

1.2. Measuring Cognitive Affective States

Even though many of the tools discussed above are static and considered less reliable [37], they continue to be popular in real-world contexts, partly because single measures are easy to administer whereas other objective measures of mental effort and affect may require expensive and hard-to-use instrumentation. By contrast, Mayer et al. [38] have highlighted the need for a direct measure of cognitive overload. While the exact nature and interplay between cognitive and affective states has been a driving force of recent research, the exact dynamics of which are still in the nascent stages, the role and importance of these cognitive-affective states in learning remains largely undisputed. There have been several encouraging attempts in multi-modal sensing of student states during learner interactions. Various technologies, such as facial expression recognition, physiological sensing such as skin conductance, heart rate variability, body temperature, posture, and neurological markers such as functional near-infrared spectroscopy (fNIRS), have been adopted to infer such cognitive-affective states. According to Calvo and D'Mello [41], the value of each modality depends on the validity of the signal as a natural way to identify an affective state, the reliability of the signals in real-world environments, the time resolutions of the signal as it relates to the specific needs of the application and the cost and intrusiveness for the user. While it is beyond the scope of the paper to provide an extensive review of each of the modalities, there are some common threads that tie most of these attempts together.

- 1. *Context of research*: Since such multi-modal sensing is central to intelligent and affect-aware tutoring systems, much of the work is conducted in specific settings, such as with intelligent tutors and learning in computer-based environments [5,42–47]. Some other research has used more experimental and laboratory settings to explore the physiological activations under stress and cognitive load (for example, [48–50]. These studies are indeed important to tease apart the underlying physiological phenomena when a person engages in different levels of mental effort, by eliminating a lot of confounding variables and the challenges that come with ambulatory settings. However, it is challenging to measure day-to-day phenomena in the sterile environment of a fully controlled laboratory, especially with fixed stimuli [51].
- 2. *Age group*: Most of the work has focused on learning in high school and university students or adults [5,8,42,47,52,53]. One of the reasons why affect has not been explored much in children is because it is hard to measure. While it is more convenient to get performance scores and test the ability to transfer learning, it is harder to measure how the learner feels during the entire process [25].

Multi-Modal Sensing towards Detection of Cognitive-Affective States

One of the popular advocacies towards employing these approaches in combination has also been called multi-modal sensing. The underlying tenet of such multi-modal approaches is the belief that emotion is an embodiment that is best captured through multiple physiological and behavioral approaches. For example, anger is expected to be manifested via particular facial, vocal, and bodily expressions and changes in physiology such as increased heart rate, and is accompanied by instrumental action [54]. Therefore, combining different responses in time will provide a better estimate of the emotion that may have ensued [55]. In spite of this, there exist few attempts with affect detection systems that integrate information from different modalities [56]. Few of the researchers have employed a combination of methods to sense affect [41,47,57]. Collecting multi-modal information in natural settings is ridden with several noise-inducing and confounding variables. Nevertheless, the advantage of multi-modal human–computer interaction systems has been recognized and seen as the next step for the mostly uni-modal approaches currently used.

With preschoolers, researchers have explored galvanic skin response as objective indicators of anxiety [58] or aggression [59]. Such work shows EDA, and specifically Skin Conductance Responses

(SCRs), as a potential physiological marker for different behaviors. EDA has been implicated in attention allocation as well [60]. Cognitive load has also been shown to have an effect on various components of Heart Rate Variability (HRV), such as mean heart rate (HR), respiratory rate, low frequency power (LF) and high frequency power (HF) components of HRV [61–63]. People under high mental workload have reduced HF components [62]. The high frequency (HF) component of heart rate variability is indicative of the parasympathetic influence on the heart and is high during rest. During high-attention tasks and tasks that elicit higher cognitive load, absolute measures of low frequency heart rate variability (LF HRV) power and LF–HF ratio have been observed to decrease when compared to a baseline [63]. Mc Duff et al. [48] used remote HRV measurements to monitor the effect of cognitive workload on heart rate variability and identified the low frequency and high frequency components of heart rate variability to be the most indicative of cognitive stress. Changes in electro-dermal activity and heart rate variability can also be mapped to other phenomena, such as changes in emotional states [64].

Kapoor et al. [65] attempted to build a system that detects surface level affect behavior, such as posture, eye-gaze, facial expressions and head movements using pressure sensors and gaze-tracking. Affect has been manually coded from observations through different coding systems. In a study conducted by Craig et al. [57], six affective states, namely frustration, boredom, flow/interest, confusion, eureka and neutral were coded. D'Mello et al. [5] used a similar descriptive system but used delight and surprise instead of eureka. The challenges with using facial expressions and speech/prosody are that they are not always discriminatory for some emotions (e.g., frustration, boredom and confusion [54] can be controlled and hardly ever occur in isolation).

1.3. Pedagogical Approaches

While we understand there are several methods of learning, we focus on two popular pedagogies that are practiced in the regional kindergartens where the study was conducted.

Constructionist Approaches: Constructionism refers to educational practices that are student-focused, meaning-based, process-oriented, interactive, and responsive to student interest [9]. According to Papert [66], "Constructionism—the N word as opposed to the V word, shares constructivism's view of learning as 'building knowledge structures' through progressive internalization of actions [...] It then adds the idea that this happens especially felicitously in a context where the learner is consciously engaged in constructing a public entity, whether it's a sand castle on the beach or a theory of the universe" (p. 2). Constructionism emphasizes knowledge construction that takes place when learners are engaged in building objects. Proponents of constructionism believe that working on context-specific tasks, which require the manipulation of virtual or concrete objects, will assist students in the learning process [67]. Further to this, Penner (2001) [68] suggests "there is a continuous interaction between thought and action. From this perspective, the starting point of all learning resides in the premise that the mind and body are extended and transformed by artefacts situated in activities" (p. 8). In our study, we follow the fundamental tenets of constructionist philosophy as applied in learning science, technology and mathematics concepts, namely, (a) identify and formulate a problem, (b) design a solution, (c) create and test a solution, (d) optimize and re-design (as needed) and (e) communicate and disseminate the solution [69].

Instructionist Approaches: In current educational contexts, instructionism refers to teacher-centered, teacher-controlled, outcome-driven, highly structured non-interactive instructional practices [9]. It is also described as systematic teaching, explicit teaching, direct teaching and active teaching [10] with high emphasis on the teacher rather than the student [11]. Teachers are often thought of as being "transmitters of objective reality" and students are viewed as "passive receptors of knowledge" [70]. There is a lot of emphasis on outlining detailed lesson plans, organization and management and teacher communication and effectiveness [71]. It therefore follows that student achievements and failures are attributed to teacher and instruction characteristics [71]. This is a commonly prevalent form of learning in several countries across age groups. Direct Instruction that was initially started by Siegfried Engelmann [72] for early literacy has now evolved to include educational practices that

7 of 21

emphasize well-developed and carefully planned lessons designed around small learning increments and clearly defined and prescribed teaching tasks [73]. In our study, we used a combination of explicit teaching, introduction of concepts with demonstrations and videos.

2. Method

2.1. Participants and Stimuli

Twelve kindergartners between aged 5 to 7 years participated in this study. The children were recruited from a kindergarten in a middle-class neighborhood. The average level of parental education was a university degree. The language of instruction in the Kindergarten was English. All the children had exposure to STEM in their class. Prior to the study, ethical approval was obtained from the Institutional Review Board (IRB). The teachers and parents of the participants were told about the aims of the research and signed parental consent, giving consent to the use of their data that was obtained. Prior to the study, the researcher spent a week immersing herself in the children's classrooms during art and craft or play time in the presence of the teacher. This was done with the purpose of building rapport and familiarity before having pull-out sessions with the children so as to make them comfortable and also eliminate any confounding effect of a stranger on data collection. Only those participants with consent and who felt comfortable were involved in the study. Every child was informed by the researcher that they can excuse themselves from the study at any time and express freely. The session itself was conducted in a room with a one-way mirror that offered views into the study session to the teachers if they chose to observe.

The overarching topic for STEM learning for the study explored the topic of *wind energy*. The materials were chosen depending on the pedagogical approach adopted as below:

Constructionist Approach: The LEGO Education's "Early Simple Machines" from their Machines and Mechanisms sets (https://education.lego.com/en-us/products/early-simple-machines-set/9656) were used. These sets are completely manual and did not require any electronic devices to get started with them. The kit (see Figure 1) included student-ready resources, such as bricks and construction guides, and also potentially useful lesson ideas and assessment tools for teachers. The teacher guides followed a constructionist approach to learning through a 4C approach of a Connect-Construct-Contemplate-Continue cycle (described in the procedure) and used this as a reference in planning the constructionist sessions. They did not require any prior science training. The curriculum highlights centered around science (concepts such as energy, force, speed, float; fair testing, predicting, collecting data), design and technology (investigating, designing, making, and testing and evaluating) and mathematics (distance, time, weight, reading scales, counting, calculating, shape and problem-solving). In addition to the resources from the kit, we used some additional materials like a fan and pieces of paper.

Instructionist Approach: The materials used for the instructionist approach included teacher explanation notes of concepts, videos demonstrating the concepts and demonstrations by the experimenter. The session followed the initiation/teaching and response cycle with the former focused on the teacher/experimenter delivering concepts through mini lecture/demo/videos and the latter focused on the child responding to questions raised by the experimenter and clarifying doubts on the part of the child if any.



Figure 1. Snapshots of the lego kit with instruction cards used for the constructionist approach.

2.2. Procedure

Every participant was introduced to the concept of wind energy through constructionist and instructionist approaches with a gap of 2 weeks between each approach (see Figure 2). The presentation of the approaches itself was counterbalanced across participants. For each approach, at the start of every session, the baseline for physiological data was collected, where each participant was asked to close their eyes and relax for 2 min. Following this, each participant learnt the concept of wind energy through the constructionist approach, each lasting about 30–40 min per constructionist session depending on the pace of the child, and learnt the concept through the instructionist approach for about 30–40 min per lesson. The sessions were counterbalanced. The structure of the session was determined by the pedagogy adopted and was constant for all the participants as outlined below:



Counter-balanced - 6 participants received constructionist approach first and the other 6 participants received instructionist approach first.

Figure 2. Presentation of pedagogical approaches to the participants.

Constructionist Approach: The sessions consisted of following phases in exact order.

- *Connect (5 min)*: The goal of this stage was to provide a context/problem scenario that presents an opportunity to help identify the problem and investigate how best to come up with a solution. The experimenter narrated a story to set a scene for participants. The experimenter then encouraged participants to come up with a solution. Following this, the participants were pointed to the resources at hand and encouraged to "build" the solution. The participant's role here was to mainly attend to the experimenter to understand the challenge and appreciate that the solution is linked to a problem.
- *Construct (10–15 min)*: The goal was to use the building instructions and build models embodying concepts related to the key learning areas. The participants were encouraged to use the cards that showed a step-by-step procedure on how to construct a pinwheel using the blocks (see Figure 1). The facilitator provided the instruction sheet and blocks and encouraged the participant to follow

the sheet and build the desired model. Further, the instructions supported the development of technical knowledge and understanding. The participant built the model following instructions and had the opportunity to raise any doubts.

- *Contemplate (10 min)*: The goal of this stage was to encourage the participants to conduct scientific investigations with their constructions. Through their investigations participants learnt to identify and compare test results. The experimenter provided questions that encouraged participants to compare two solutions and reason their answers (e.g., "Predict which of the wheels will start turning near the fan" or, "How far from the fan will the pinwheel start turning"). The participant was asked to reflect on what happened, why it happened, if it matched their predictions, how the model works, if the test was fair and in general describe what happened and to reflect and share their thoughts on questions posed by the experimenter.
- *Question and answer (5 min)*: At the end of the session, the participants were asked five questions that tested the concepts learnt. These included multiple choice, open-ended and true/false questions.

Instructionist Approach: The instructionist session had the following three elements spread throughout the session similar to the way the lessons were typically carried out in the school where data were collected. The question and answer phase was always at the end as it tested an understanding of the concepts using the same questions as in the constructionist approach.

- *Teaching concept* (10–15 *min*): The experimenter introduced the concept through a lecture where energy and wind energy were explained with definition and examples. The content taught was the same as the constructionist approach, except it took the form of a lecture.
- *Videos (8–10 min)*: The participants viewed two videos that demonstrated the concept of energy and its forms and wind energy (what it means and uses with examples).
- *Demonstration (10 min)*: The teacher finally demonstrated wind energy using an electric fan and manually blowing different materials such as leaves, paper, foil, tissue, ball, chair, toy etc.
- *Question and answer (5 min)*: At the end of the session, the participants were asked five questions that tested the concepts learnt. These included multiple choice, open-ended and true/false questions.

2.3. Data Collection

Each participant wore an E4 wristband (https://www.empatica.com/research/e4/) throughout the session, this recorded electrodermal activity and heart rate measures. The sessions were videotaped to later code for facial expressions, eye gaze, posture, head movements, verbal expressions and any other prominent non-verbal behavior.

2.4. Data Analysis

Physiological measures: We extracted two measures of skin conductance across the tasks: (a) the number of Skin Conductance Responses (SCRs) and (b) average amplitude of the SCRs. Any change in skin conductance greater than 0.01 microsiemens (μ S) was considered as an SCR [74]. A continuous decomposition analysis was done using the Ledalab toolkit for Matlab 2016b [75]. We analyzed various measures using Kubios HRV 2.2 [76] on Matlab 2016b. The analyzed measures included: mean Heart Rate (HR), mean inter-beat-intervals (RR), Heart Rate Variability (HRV) and Low Frequency (LF) and High Frequency (HF) components of heart rate variability.

Observational Coding: In this study, we used an affect coding index as a reference for coding affect along with the coding of other non-verbal and verbal behavior as used by Sridhar et al. [6] and outlined in Figure 3. This included facial expressions, eye contact, head movements and posture as well as six affective states, namely delight, confusion, frustration, boredom, flow and neutral. Two coders first independently coded the video recordings after the sessions. Following this, they discussed the coded behaviors to arrive at a consensus. This was done as the intention was to understand the evolution of cognitive affective states across the stages rather than establishing the inter-rater reliability between the two coders.

Performance Scores: For the five questions that were asked as part of the Q&A phase, every correct response was scored 1 and a partially correct response (for open-ended questions) was scored 0.5. Therefore, every participant was scored within a maximum score of 5 points.

Affective State	Definition	Signs
Neutral	No apparent emotion or feeling	A void of emotion and no facial features or if no emotions could be determined
		Did not appear to be displaying any of the affective states below, or the student's affect could not be determined for certain.
Boredom	Being weary or restless through lack of interest	Participants seemed uninterested in the activity or responded slowly to the system and did not appear motivated.
		Behaviours such as slouching, and resting the chin on his/her palm; statements such as "Can we do something else?" and "This is boring!"
Frustration	Dissatisfaction or annoyance	behaviors such as banging on the screen/ device or pulling at his/her hair/ running hands through hair; statements such as "This is annoying!" or "What's going on?!?"
Confusion	A noticeable lack of understanding	Participants seemed puzzled and not sure how to continue or were struggling to understand the material.
		Behaviors such as scratching his/her head, repeatedly looking at the same interface elements; statements such as "Why didn't it work?"
Delight	A high degree of satisfaction	Behaviors such as clapping hands or laughing with pleasure; statements such as "Yes!" or "I got it!"
Surprise	Taken aback/ encounters something totally unexpected	Behaviors such as jerking back suddenly or gasping; statements such as "Huh?" or "Oh, no!" $% \label{eq:state}$
Eureka	Sudden switches from seemingly negative to positive state (not to be mistaken with undesirable to desirable)	If participant is observed to transfer from a state of confusion to a state of intense interest, as manifested by responding very quickly after a period of inactivity.
Flow/interest/engaged	State of interest that results from involvement in an activity	When participants showed interest in the interaction or were paying attention and responding quickly.
		Behaviors such as leaning towards the device/activity or mouthing solutions to him/herself while solving a problem

Figure 3. Affect Coding Index.

3. Results and Discussion

The findings from each of the data sources are presented for each pedagogy and discussed through a presentation of overall trends and case illustrations.

3.1. Constructionist Approach

Physiological, observational and performance data were analyzed to understand what participants went through as they learnt the concept of wind energy across the connect, construct and contemplate phases. The construct (M = 70, SD = 73) and contemplate (M = 64, SD = 47) phases had a significantly higher number of SCRs compared to the connect phase (M = 11, SD = 11), as revealed by a one way ANOVA (F = 4.96, p < 0.01). Connect was a more passive stage that required the participant to only listen to the experimenter without actively participating. However, the construct and contemplate phases were more active as they required the child to follow instructions on the card to construct a pinwheel and then reason about why their construction worked or failed, accordingly. This elicited higher attention and engagement as demonstrated through increased SCRs. The same trend was seen with mental effort as revealed by the analysis of heart rate variability as well. The participants expended significantly more mental effort (F = 4.26, p < 0.05) in the construct (M = 0.87, SD = 0.42) and contemplate (M = 0.61, SD = 0.44) phases compared to the connect phase (M = 0.41, SD = 0.29). Again, this may be attributed to the nature of the phase itself.

While physiological findings pointed to a more underlying response to learning, the observational coding pointed to overt manifestations of participants' expressed behavior as they learnt the concept across the phases. Since every phase contributed to the learning of the concept of wind energy and elicited different responses and engaged different mental faculties, we did not expect to see an increase in one type of affect over another but rather see how each of these elicited different affective responses.

We identified the top three affects for each phase (see Figure 4) in understanding the most dominant affect types. The connect phase was primarily characterized by two affective states: neutral and flow. As participants listened to the experimenter they usually seemed attentive and had good eye contact, sometimes leaning in out of curiosity to listen to the story but not necessarily displaying any other overt signs of other affect. They were only required to listen to a story that was extremely simple and easy to understand, which perhaps meant that there was not much scope to be confused or bored. The construct phase saw participants being delighted as they figured out constructions ("Where does this one go... oh... got it!"), confused when they could not assemble a part or could not understand why it did not fit the same in spite of them thinking they were following the visual instruction card, frustrated when they were confused for a long time without finding a solution and sometimes just calm without displaying any overt signs. Flow was the most dominant affect as participants were absorbed in building their pinwheels. Here again, the participants maintained good eye contact and posture, only occasionally lifting their head up to seek help. The instances of negative affect were fewer in that there was always a visual instruction sheet that showed how to construct the pinwheel. The contemplate phase had the richest variety of affects, as participants displayed all of the six affective behaviors given the nature of the phase, where children were asked to test their construction fairly and reason when it would work best. The most dominant negative affects were frustration and confusion when they could not figure out a solution or reasoning. On the other hand, they also demonstrated delight when their pinwheels worked. In spite of the diversity in how different participants reacted, the performance of the participants did not reveal vast differences (M = 4.0; SD = 0.69) and all the participants scored above 3 out of a maximum of 5.



Figure 4. Stages of constructionist and instructionist approaches.

Unlike the constructionist approach, the instructionist approach did not follow a phase-wise progression. In order to parallel the analysis of the constructionist approach and facilitate as fair a comparison as possible, we analyzed the data according to the key tasks: (1) demonstrations by the instructor, (2) teaching of a concept and (3) watching videos. The results are shown in Figure 4. There was no statistically significant difference in terms of engagement as revealed by a one-way ANOVA (F = 0.81, p > 0.05) comparing the number of SCRs across these phases. This may have been due to large individual differences within each condition. For example, P11 showed the largest number of SCRs for watching the demonstration as compared to other conditions, P6 had very similar SCRs for all conditions while P8 had the largest number of SCRs for the teaching concept phase and lowest SCRs for watching the demonstration and videos phases. Similarly, no significant difference was found in terms of mental effort expended across the conditions (F = 1.84, p > 0.05).

Unlike the constructionist approach, there were two affective states that dominated the phases, namely neutral and flow. The demonstration phase was predominantly characterized by flow as the experimenter demonstrated the concept of wind energy while the participant watched. Watching videos was characterized mainly as neutral, as participants did not display any overt signs or demonstrate increased flow.

Across the three phases, flow seemed to be highest when a teacher was involved, as seen by increased flow in the demonstration and teaching concept phases. The limited range of affects may be attributed to the nature of the phases and type of participation elicited from the learner. Even though participants seemed alert and exhibited good eye contact and attention span, they functioned more as information recipients in contrast to the constructionist approach where participants had to actively demonstrate learning throughout except the connect phase. Interestingly, the connect phase mimics the instructionist approach in that it was characterized by neutral and flow affects as well. However, performance scores did not reveal vast differences (M = 3.8; SD = 0.45) and all the participants scored above 3 out a maximum of 5.

3.3. Comparison of the Two Approaches

Although there was no significant difference in the performance score (t(11) = 1.30, p = 0.11), there were overarching patterns across the constructionist and instructionist pedagogies (see Figure 5) as discussed below.

Range of affect: Clearly the range of affect was higher in the constructionists approach, with participants demonstrating all the 6 affective states across the phases (see Figure 5 *Affect*). By contrast, the instructionist approach was marked by neutral or flow states. In the constructionist approach, there was a lot of movement between states that can be likened to moving in and out of quadrants as proposed by Kort et al. [19] This may have been owing to the nature of the task itself, where children did most of the problem-solving with the facilitator only helping them as needed. It has been shown that certain negative states, such as frustration and confusion, are indeed beneficial to learning better [15,77], and we found the presence of such states more in the constructionist approach compared to the instructionist approach. Though the argument that pedagogies that challenge and frustrate may seem counterintuitive, it is true that learning environments need to challenge the learner in order to elicit critical thought and deep inquiry [5], and this may be absent in shallow delivery approaches.

Patterns of arousal: The participants were in a significantly lower state of arousal (t(35) = -2.664, p < 0.01) in the instructionist approach compared to the constructionist approach (see Figure 5 *EDA*). This may have been due to the nature of the tasks in the instructionist approach, which required lesser participation from the participants. Incidentally, the connect phase within the constructionist approach had the lowest number of SCRs, where again the participants' role was to only listen to the experimenter. Therefore, the relatively reduced arousal or engagement may be viewed as a reflection of the nature of the phase itself. While reduced SCRs may also mean that learners were comfortable with the learning tasks, such environments when extended for long periods of time do not cause learning, as there

are no challenges to inspire a deep enquiry. However, since this was only one learning session, it is challenging to remark on the learning outcomes or attribute it to the presence of such states alone.

Patterns of mental effort: The heart rate measurements also indicated a higher mental effort in more phases of the constructionist approach, and particularly the construct and contemplate phases as compared to the instructionist approach (see Figure 5 *HRV*). The higher LF/HF value seemed to correspond with a negative affect or challenge in constructionist approach. This also mimicked the patterns with SCRs, suggesting possible problem-solving or thinking, especially in the construct and the contemplate phases of the constructionist approach. By contrast, participants were more relaxed (not significant though, t(35) = -0.93, p = 0.18) in the instructionist approach, as shown by the lower LF/HF especially while viewing videos. The highest effort in the instructionist approach mapped to those phases where an experimenter was involved in delivering information to the participant that may have caused an increased effort to process what was being taught. This is also in tandem with the findings from the analysis of number of SCRs.



Figure 5. Comparison of constructionist and instructionist approaches.

3.4. Case Illustrations

In spite of several overall overarching commonalities that emerged across the pedagogies, there were individual differences in how participants responded to learning across the phases. Individual learning styles and preferences will always influence how learners respond to pedagogies and this was no different. We present two case illustrations below to highlight some of such nuances that emerged as we analyzed the data sources across individuals and phases.

Consider P2 (see Figure 6a), who had similar scores of 4 and 3.5 in the instructionist and constructionist approaches, respectively. In the connect phase of the constructionist approach, he had the lowest arousal and mental effort as he only listened to the experimenter. Once he moved to the construct phase and was expected to build the pinwheel, the arousal and accompanying mental effort increased and this corresponded with the dominant affect patterns of frustration and delight as he figured out how to build the pinwheel. As he moved to the contemplate phase and had to actively test his pinwheel, the arousal increased further and the mental effort dropped but remained moderate. The participant demonstrated instances of frustration that may have contributed to the mental effort and arousal but otherwise remained neutral. Now let us consider his response in the instructionist approach. The highest arousal and effort were seen as the participant watched the demonstration mostly in a state of flow. The videos and the teaching concept phase were both marked by lower arousal and effort without any other prominent affect. One can draw strong parallels within the participant's responses across the connect, videos and the teaching concept phases, all of which involved passive participation from the participant. Even though the demonstration was done by the experimenter, witnessing a concept in the form of "show and tell" may have been interesting to P2, as the arousal and mental effort pattern mimicked the construct and contemplate phases. Thus for P2, a more hands-on approach to learning or including demonstrations instead of lectures or audio-visual aids. materials would probably be a more effective way to engage the learner and build a good challenge.

Consider P11 (see Figure 6b) who also scored similar scores of 4.5 and 5 in the instructionist and constructionist approaches, respectively. He had very high levels of arousal (demonstration being

highest) with approximately similar levels of effort across all the phases in the instructionist approach. The demonstration and teaching of concept, both of which involved the presence of an experimenter, were marked by the affective state of flow, unlike the watching of videos, where it remained largely neutral. However, in the constructionist approach, the arousal levels were very low compared to the instructionist approach. This was also different from the overall findings, where the arousal was significantly higher on average for the constructionist approach. The effort was highest for the construct phase but similarly low for the connect and contemplate phases. The participant had instances of boredom in the contemplate phase that may suggest the activity to have been very easy or tiring but the maximum score of 5 and the observations during the contemplate phase point to the activity being easy for the learner. In this instance one might argue that a challenging concept may be introduced through a hands-on phase-wise approach and through a more direct face-to-face interaction with the teacher since they elicited a flow state for the most part and also had the learner most engaged as seen through the SCRs. Of course, such pedagogical decisions must be driven by more similar observations and long-term studies. Still, appreciating that such nuanced differences can indeed be teased by looking at all the data comprehensively and used to guide learning is encouraging. Our conversations with teachers indeed reinforced this belief. While these case illustrations only offer a sneak peak into individual learning behaviors across two pedagogies, more such detailed studies in a classroom context performed across concepts and related contexts may help identify individual learning styles, responses to newer forms of teaching and learning interfaces themselves.



Figure 6. Case illustrations.

4. Recommendations for Running Similar Studies

Based on our experience from running the study and previous work such as [78], we have synthesized specific considerations for researchers performing similar work with children. We have categorized these recommendations into general, EDA-related and HRV-related considerations as follows:

4.1. General Considerations

- i *Importance of establishing rapport*: Any procedure with children as participants, especially those that look at emotional/physiological responses, must begin with rapport building and data must be collected from a very familiar environment. In this study, the experimenter spent a week with the participants in their classroom in art and play activities.
- ii *Including repeated baselines and breaks between tasks*: If the participants undergo different activities, it is recommended that a baseline be established before every task to estimate changes in physiological measures. It is also recommended that the entire procedure be conducted under controlled ambient conditions, as room temperature and humidity may influence the physiological and to some extend emotional responses.
- iii *Familiarity with on-seat tasks and compliance*: Since this was a preschool group who were more familiar with compliance and on-seat behaviors, there were not many issues with compliance with experimental instructions, which is an important factor for learning tasks.
- iv *Effect on task time on physiological measurements*: Time taken to complete a task has an effect on the measurements. For example, a very short task may not capture the fluctuation in the physiological parameters well. A very long task, on the other hand, may bore or tire the child, thereby discouraging him from even continuing participation. Using a combination of short and longer tasks may best mimic learning in real-life situations. Tasks can also be categorized as being consistently easy at the start and increasing in difficulty towards later parts, or have a mix of easy and difficult items interspersed. Having mixed-difficulty tasks may give a good insight into whether the measures are truly responsive to randomly occurring difficulties and cognitive load and not just built up over time.
- v *Keeping instructions simple and clear with ample opportunities to test understanding*: It is recommended to have ample opportunities for the practice task. This is to ensure that the performance on the tasks and other physiological data collected from the task are a reflection of the ability tested by the task rather than errors resulting from misinterpretation or lack of understanding.

Specific Considerations in Eliciting EDA Data

- i *Capturing non-specific SCR (NS-SCRs) for learning tasks*: NS-SCRs are those responses that are not associated with discrete stimuli. Such responses are not measured in terms of amplitude but rather as the number of SCRs per minute or over the time period of activity. A minimum value must be specified as the threshold. Current sensors allow for setting thresholds as low as 0.01 microsiemens, which is more appropriate for small children, as the NS-SCRs with no sudden external stimuli do not evoke the same intensity of response as in the case of sudden stimuli.
- ii Monitoring any drastic changes in the tonic component of SCR: According to Dawson et al., (2011) [60], increases in non-SCRs are attributed to (a) an increase in tonic arousal, energy regulation or mobilization, (b) attentional and information processing, or (c) stress and affect. The tonic arousal can be ruled out by looking at any abrupt increases in Skin Conductance Levels (SCLs) over time windows.
- iii *Awareness of developmental effects on SCRs*: SCRs in general are reported to appear later in children and even though they develop fairly well by 5–6 years [78], as we noticed in the responses in the study, it is recommended to include video taping of the sessions in order to capture other signs when physiological signals are missing. Further, since the SCR does not offer data on valence, the video-taping will facilitate interpretation of SCRs to some extent.

4.2. Specific Considerations in Eliciting HRV Data

- i *Accounting for data loss from insufficient contact between photoplethysmography (PPG) sensor sensor and small wrist size*: In spite of the adjustable strap and getting the tightest fit, the PPG sensor for heart rate measurement sometimes did not achieve good contact with the wrist of the participant owing to their small wrist size, which resulted in data loss. In order to prevent this, we used a small pad of cloth near the strap to enable better contact of the PPG sensor on the wrist, especially for children of smaller build.
- ii Eliminating artifacts and accounting for data loss from optical noise: The biggest challenge with HR measurements from optical sensors is missing data from motion noise. As a result, there were periods where the HR measurements were not recorded because the child moved their hands a lot. While small movements do not affect this, sudden big movements result in a loss of data. Therefore if the child is exerting mental effort but makes a movement during this period, there is always a risk of losing the data. Therefore using video data and other contextual cues is important to decipher why data are missing or what may have ensued during this period. It is recommended to check for noise when children make general normal motor movements like moving of hands, head movements etc. when performing an activity. In our case, we did not find this to impact their skin conductance data, while some movements resulted in loss of heart rate measures for that brief period. Being aware of how such movement impact data is important when interpreting the data. For example, it helps to look at the accelerometer data to see if there was a movement if video tapes do not present very obvious movements.

5. Limitations

As observed by Calvo et al. [5] no single "sophisticated synchronized response that incorporates peripheral physiology, facial expression, speech, modulations of posture, affective speech, and instrumental action" (p. 36) emerged for every affect. Though there seemed to be some uniform patterns in certain sources of data for all participants, such as increased number of SCRs for delight and surprise (in the instances observed) accompanied by a low LF/HF value, such an integrated multi-modal pattern across data channels was not seen for other affect behaviors such as confusion and frustration. This may have been due to some ambiguity in coding such affects or the absence of a synchronized set of responses that differentiate these two affect behaviors given the data channels considered. Within skin conductance, there were indeed some commonalities for certain categories of emotion that seemed to be shared by most participants. For example, delight, confusion and frustration were often accompanied by high arousal as observed by Woolf et al. [8], and looking away or lack of eye contact seemed to have a plateau or decrease in skin conductance. There existed no such trends with heart rate measures, as they are more of a proxy of cognitive load than emotion alone [25,48]. The study did not aim to uncover universal patterns in data corresponding to an affect but rather to use the physiological measures as a means to complement and supplement overt observational data or lack thereof [47], as illustrated through some cases.

Despite analyzing the coded behaviors in quadrants for each phase, we did not find certain states to dictate certain quadrants as shared by Kort et al. [19] but rather several jumps between them. This may have been due to the nature of one-on-one sessions, the duration of the task or the age group. Further, none of the states prevailed for a long duration but it may be possible that certain states, such as boredom, may dictate the entire session when in a classroom [77].

We did not employ pre- and post-assessment measures, as we did not aim to capture any change in knowledge. Additionally, learning is very difficult to assess except in cases such as remembering [79]. The small sample size and the specific social and educational environment does not allow for generalization. However, employing such assessments over longer periods of learning may offer interesting insights into the role of cognitive affective states and transitions from one state to another, as well as learning gains and how they can supplement teaching in classrooms. Examining this in light of individual learning styles and preferences may help form more generalizable

inferences accordingly. Conversely, the case illustrations do offer a potential way of supplementing understanding of such preferences themselves. While the method and analyses have looked at pedagogical approaches separately, we are aware that teaching evolves organically in a classroom, where the teacher chooses from an array of approaches that best meet the context. There is a high likelihood that a classroom will be hybrid in terms of the pedagogical approaches. In such an instance, we believe that looking at the differences in the phases and how they elicited different cognitive affective states may offer support or suggestions for the choice of pedagogies. To emphasize, we do not propose that pedagogies and teaching be dictated by the findings of combining such data sources alone, but rather more as a reflective process for educators, designers and researchers to use as a potential tool in understanding what may have ensued and contributed to an observed performance.

6. Conclusions and Future Work

In the paper, we conducted a study with 12 kindergartners to explore the role of physiological, behavioral and observational data to understand cognitive-affective states as they learnt the concept of wind energy through constructionist and instructionist approaches. We found that triangulation of these data channels is especially useful in tracking an individual's learning process and also offers a potential way to compare pedagogies by way of exploring preferences and underlying cognitive-affective states.

As models of affect and affective sensing refine over the years, we foresee triangulation between different and relevant modalities of data to play a pivotal role in continuing to supplement and unearth covert mechanisms. Research in educational technology must move beyond researchers' perspectives and into the real world, where teachers can understand and use such technology to explore students' learning and behavior [80]. Our ongoing research plans include analyzing and mapping the responses of the children during learning in the current study, and testing such methods within a wider range of populations across different learning contexts and technologies in natural and pull-out settings. While deploying such measures in classrooms may seem challenging, the field of affective learning has indeed come far with wearable sensors, analysis methods and our encouraging conversations with teachers who find value in such measures in light of their current needs. All of this only reinforces our belief that such methods will pave the way for designing pedagogies, learning tools and other adaptive learning interfaces that are responsive to the learner's cognitive and affective states.

Author Contributions: P.K.S. conceptualised and conducted the investigation as part of her PhD research. S.N. contributed with overall supervision, review editing. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Assistive Augmentation research grant under the Entrepreneurial Universities (EU) initiative of New Zealand.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. Edith, A. *Piaget's Constructivism, Papert's Constructionism: What's the Difference;* Massachusetts Institute of Technology: Cambridge, MA, USA, 2002.
- 2. Resnick, L.B.; Resnick, D.P. Assessing the thinking curriculum: New tools for educational reform. In *Changing Assessments*; Springer Netherlands: Dordrecht, The Netherlands 1992; pp. 37–75.
- 3. Ontario Ministry of Education. *Growing Success: Assessment, Evaluation and Reporting in Ontario Schools;* Ontario Ministry of Education: Toronto, ON, Canada, 2010.
- 4. Pekrun, R. The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educ. Psychol. Rev.* **2006**, *18*, 315–341. [CrossRef]
- D'Mello, S.; Graesser, A. Dynamics of affective states during complex learning. *Learn. Instr.* 2012, 22, 145–157. [CrossRef]

- Sridhar, P.K.; Chan, S.W.; Nanayakkara, S. Going beyond performance scores: Understanding cognitive-affective states in kindergarteners. In Proceedings of the 17th ACM Conference on Interaction Design and Children, Trondheim, The Netherlands, 19–22 June 2018; pp. 253–265.
- 7. O'Donoghue, T.; Punch, K. *Qualitative Educational Research in Action: Doing and Reflecting*; Routledge: Abingdon, UK, 2003.
- 8. Woolf, B.; Burleson, W.; Arroyo, I.; Dragon, T.; Cooper, D.; Picard, R. Affect-aware tutors: Recognising and responding to student affect. *Int. J. Learn. Technol.* **2009**, *4*, 129–164. [CrossRef]
- 9. Johnson, G.M. Instructionism and Constructivism: Reconciling Two Very Good Ideas. Int. J. Spec. Educ. 2005, 24, 90–98.
- Schug, M.C.; Tarver, S.G.; Western, R.D. Direct Instruction and the Teaching of Early Reading: Wisconsin's Teacher-Led Insurgency; Wisconsin Policy Research Institute Report; Wisconsin Policy Research Institute: Thiensville, WI, USA, 2001; Volume 14, Number 2.
- 11. Jonassen, D.H. *Computers in the Classroom: Mindtools for Critical Thinking*; Prentice-Hall, Inc: Upper Saddle River, NY, USA, 1996.
- 12. Sumida, M. Kids science academy: Talent development in STEM from the early childhood years. In *Science Education in East Asia*; Springer International Publishing: Cham, The Netherlands 2015; pp. 269–295.
- 13. Jones, A.L.; Stapleton, M.K. 1.2 million kids and counting—Mobile science laboratories drive student interest in STEM. *PLoS Biol.* **2017**, *15*, e2001692.
- 14. Kessen, W.; Ortony, A.; Craik, F. *Memories, Thoughts, and Emotions: Essays in Honor of George Mandler;* Psychology Press: Hove, UK, 2013.
- 15. Dweck, C.S. Motivational processes affecting learning. Am. Psychol. 1986, 41, 1040–1048. [CrossRef]
- 16. Ames, C.; Archer, J. Achievement goals in the classroom: Students' learning strategies and motivation processes. *J. Educ. Psychol.* **1988**, *80*, 260–267. [CrossRef]
- 17. Dweck, C.S.; Leggett, E.L. A social-cognitive approach to motivation and personality. *Psychol. Rev.* **1988**, *95*, 256–273. [CrossRef]
- 18. Csikszentmihalyi, M. Finding Flow; Basic Books: New York, NY, USA, 1997.
- Kort, B.; Reilly, R.; Picard, R.W. An affective model of interplay between emotions and learning: Reengineering educational pedagogy-building a learning companion. In Proceedings of the IEEE International Conference on Advanced Learning Technologies, Madison, WI, USA, 6–8 August 2001; pp. 43–46.
- Pekrun, R.; Goetz, T.; Daniels, L.M.; Stupnisky, R.H.; Perry, R.P. Boredom in achievement settings: Exploring control-value antecedents and performance outcomes of a neglected emotion. *J. Educ. Psychol.* 2010, 102, 531. [CrossRef]
- 21. Conati, C.; Maclaren, H. Empirically building and evaluating a probabilistic model of user affect. *User Model*. *User Adapt. Interact.* **2009**, *19*, 267–303. [CrossRef]
- Forbes-Riley, K.; Litman, D. Metacognition and learning in spoken dialogue computer tutoring. In Proceedings of the International Conference on Intelligent Tutoring Systems, Pittsburgh, PA, USA, 14–18 June 2010; pp. 379–388.
- Hussain, M.S.; AlZoubi, O.; Calvo, R.A.; D'Mello, S.K. Affect detection from multichannel physiology during learning sessions with AutoTutor. In Proceedings of the International Conference on Artificial Intelligence in Education, Auckland, New Zealand, 28 June–1 July 2011; pp. 131–138.
- 24. Bransford, J.D.; Schwartz, D.L. Chapter 3: Rethinking transfer: A simple proposal with multiple implications. *Rev. Res. Educ.* **1999**, *24*, 61–100. [CrossRef]
- 25. Picard, R.W.; Papert, S.; Bender, W.; Blumberg, B.; Breazeal, C.; Cavallo, D.; Machover, T.; Resnick, M.; Roy, D.; Strohecker, C. Affective learning—A manifesto. *BT Technol. J.* **2004**, *22*, 253–269. [CrossRef]
- 26. Sutton, R. Assessment for Learning; RS Publications: Manchester, UK, 1995.
- Harlen, W. On the relationship between assessment for formative and summative purposes. *Assess. Learn.* 2006, 2, 95–110.
- Garrison, C.; Ehringhaus, M. Formative and summative assessments in the classroom. Available online: https://www.amle.org/browsebytopic/whatsnew/wndet/tabid/270/artmid/888/articleid/286/ formative-and-summative-assessments-in-the-classroom.aspx (accessed on 19 June 2018).
- 29. York, B.N.; Loeb, S. One Step at a Time: The Effects of an Early Literacy Text Messaging Program for Parents of *Preschoolers*; Technical Report; National Bureau of Economic Research: Cambridge, MA, USA, 2014.
- 30. Clune, W.H.; White, P.A. *Policy Effectiveness of Interim Assessments in Providence Public Schools*; WCER Working Paper No. 2008-10; Wisconsin Center for Education Research (NJ1): Madison, WI, USA, 2008.

- 31. Maxwell, G.S. Teacher observation in student assessment. Available online: https://digitised-collections.unimelb. edu.au/bitstream/handle/11343/115657/scpp-00437-qld-2001.pdf?sequence=1 (accessed on 23 June 2018).
- Read, J.C.; MacFarlane, S. Using the fun toolkit and other survey methods to gather opinions in child computer interaction. In Proceedings of the 2006 Conference on Interaction Design and Children, Tampere, Finland, 12–14 June 2006; pp. 81–88.
- 33. Pociask, F.D.; Morrison, G.R. Controlling split attention and redundancy in physical therapy instruction. *Educ. Technol. Res. Dev.* **2008**, *56*, 379–399. [CrossRef]
- 34. Hasler, B.S.; Kersten, B.; Sweller, J. Learner control, cognitive load and instructional animation. *Appl. Cogn. Psychol.* **2007**, *21*, 713–729. [CrossRef]
- 35. Yuksel, B.F.; Oleson, K.B.; Harrison, L.; Peck, E.M.; Afergan, D.; Chang, R.; Jacob, R.J. Learn piano with BACh: An adaptive learning interface that adjusts task difficulty based on brain state. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, 7–12 May 2016; pp. 5372–5384.
- 36. Leppink, J.; van den Heuvel, A. The evolution of cognitive load theory and its application to medical education. *Perspect. Med Educ.* **2015**, *4*, 119–127. [CrossRef]
- 37. Picho, K.; Artino, A.R., Jr. 7 deadly sins in educational research. J. Grad. Med. Educ. 2016, 8, 483–487. [CrossRef]
- Mayer, R.E.; Hegarty, M.; Mayer, S.; Campbell, J. When static media promote active learning: Annotated illustrations versus narrated animations in multimedia instruction. *J. Exp. Psychol. Appl.* 2005, 11, 256–265. [CrossRef]
- Nakasone, A.; Prendinger, H.; Ishizuka, M. Emotion recognition from electromyography and skin conductance. In Proceedings of the 5th International Workshop on Biosignal Interpretation, Tokyo, Japan, 6–8 September 2005; pp. 219–222.
- Lyu, Y.; Luo, X.; Zhou, J.; Yu, C.; Miao, C.; Wang, T.; Shi, Y.; Kameyama, K.I. Measuring photoplethysmogram-based stress-induced vascular response index to assess cognitive load and stress. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, Seoul, Korea, 18–23 April 2015; pp. 857–866.
- 41. Calvo, R.A.; D'Mello, S.K. *New Perspectives on Affect and Learning Technologies*; Springer Science & Business Media: Berlin, Germany, 2011; Volume 3.
- 42. Vail, P.L. Emotion: The on/off Switch for Learning; Modern Learning Press: Rosemont, NJ, USA, 1994.
- 43. Hartley, D.; Mitrovic, A. Supporting learning by opening the student model. In Proceedings of the International Conference on Intelligent Tutoring Systems, San Sebastian, Spain, 2–7 June 2002; pp. 453–462.
- 44. Gluga, R. Long term student learner modeling and curriculum mapping. In Proceedings of the International Conference on Intelligent Tutoring Systems, Pittsburgh, PA, USA, 14–18 June 2010; pp. 227–229.
- 45. Corbett, A.T.; Anderson, J.R. Student modeling and mastery learning in a computer-based programming tutor. In Proceedings of the International Conference on Intelligent Tutoring Systems, Montreal, QC, Canada, 10–12 June 1992; pp. 413–420.
- Stevens, R.; Soller, A.; Cooper, M.; Sprang, M. Modeling the development of problem solving skills in chemistry with a web-based tutor. In Proceedings of the International Conference on Intelligent Tutoring Systems, Maceio, Brazil, 30 August–3 September 2004; pp. 580–591.
- 47. D'Mello, S.K.; Graesser, A. Multimodal semi-automated affect detection from conversational cues, gross body language, and facial features. *User Model. User Adapt. Interact.* **2010**, 20, 147–187. [CrossRef]
- McDuff, D.J.; Hernandez, J.; Gontarek, S.; Picard, R.W. Cogcam: Contact-free measurement of cognitive stress during computer tasks with a digital camera. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, 7–12 May 2016; pp. 4000–4004.
- Shi, Y.; Ruiz, N.; Taib, R.; Choi, E.; Chen, F. Galvanic skin response (GSR) as an index of cognitive load. In Proceedings of the CHI'07 extended abstracts on Human factors in computing systems, San Jose, CA, USA, 28 April–3 May 2007; pp. 2651–2656.
- 50. Setz, C.; Arnrich, B.; Schumm, J.; La Marca, R.; Tröster, G.; Ehlert, U. Discriminating stress from cognitive load using a wearable EDA device. *IEEE Trans. Inf. Technol. Biomed.* **2010**, *14*, 410–417. [CrossRef]
- 51. Hedman, E.B. Thick Psychophysiology for Empathic Design. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2014.
- 52. Arroyo, I.; Cooper, D.G.; Burleson, W.; Woolf, B.P.; Muldner, K.; Christopherson, R. Emotion sensors go to school. In Proceedings of the AIED 2009, Brighton, UK, 6–10 July 2009; Volume 200, pp. 17–24.

- Vail, A.K.; Grafsgaard, J.F.; Boyer, K.E.; Wiebe, E.N.; Lester, J.C. Predicting learning from student affective response to tutor questions. In Proceedings of the International Conference on Intelligent Tutoring Systems, Zagreb, Croatia, 7–10 June 2016; pp. 154–164.
- 54. Calvo, R.A.; D'Mello, S. Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Trans. Affect. Comput.* **2010**, *1*, 18–37. [CrossRef]
- 55. Ekman, P. Facial expression and emotion. Am. Psychol. 1993, 48, 384–392. [CrossRef] [PubMed]
- 56. Jaimes, A.; Sebe, N. Multimodal human-computer interaction: A survey. *Comput. Vis. Image Underst.* **2007**, *108*, 116–134. [CrossRef]
- 57. Craig, S.; Graesser, A.; Sullins, J.; Gholson, B. Affect and learning: An exploratory look into the role of affect in learning with AutoTutor. *J. Educ. Media* **2004**, *29*, 241–250. [CrossRef]
- Najafpour, E.; Asl-Aminabadi, N.; Nuroloyuni, S.; Jamali, Z.; Shirazi, S. Can galvanic skin conductance be used as an objective indicator of children's anxiety in the dental setting? *J. Clin. Exp. Dent.* 2017, 9, e377–e383. [CrossRef] [PubMed]
- 59. Posthumus, J.A.; Böcker, K.; Raaijmakers, M.; Van Engeland, H.; Matthys, W. Heart rate and skin conductance in four-year-old children with aggressive behavior. *Biol. Psychol.* **2009**, *82*, 164–168. [CrossRef]
- 60. Dawson, M.E.; Schell, A.M.; Courtney, C.G. The skin conductance response, anticipation, and decision-making. *J. Neurosci. Psychol. Econ.* **2011**, *4*, 111–116. [CrossRef]
- Bousefsaf, F.; Maaoui, C.; Pruski, A. Remote assessment of the heart rate variability to detect mental stress. In Proceedings of the 2013 7th International Conference on Pervasive Computing Technologies for Healthcare and Workshops, Venice, Italy, 5–8 May 2013; pp. 348–351.
- 62. Hjortskov, N.; Rissén, D.; Blangsted, A.K.; Fallentin, N.; Lundberg, U.; Søgaard, K. The effect of mental stress on heart rate variability and blood pressure during computer work. *Eur. J. Appl. Physiol.* **2004**, *92*, 84–89. [CrossRef]
- Moses, Z.B.; Luecken, L.J.; Eason, J.C. Measuring task-related changes in heart rate variability. In Proceedings of the 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Lyon, France, 22–26 August 2007; pp. 644–647.
- 64. Fairclough, S.H. Fundamentals of physiological computing. Interact. Comput. 2008, 21, 133–145. [CrossRef]
- 65. Kapoor, A.; Mota, S.; Picard, R.W. Towards a learning companion that recognizes affect. In Proceedings of the AAAI Fall Symposium, North Falmouth, MA, USA, 2–4 November 2001; pp. 2–4.
- 66. Papert, S.; Harel, I. Situating constructionism. *Constructionism* **1991**, *36*, 1–11.
- 67. Carbonaro, M. Making technology an integral part of teaching: The development of a constructionist multimedia course for teacher education. *J. Technol. Teach. Educ.* **1997**, *5*, 255–280.
- 68. Penner, D.E. Chapter 1: Cognition, computers, and synthetic science: Building knowledge and meaning through modeling. *Rev. Res. Educ.* **2000**, *25*, 1–35. [CrossRef]
- 69. Rogers, C.; Portsmore, M. Bringing engineering to elementary school. J. STEM Educ. Innov. Res. 2004, 5, 17–28.
- 70. Kazdin, A.E. Behavior Modification in Applied Settings; Waveland Press: Long Grove, IL, USA, 2012.
- 71. Adams, G.L.; Engelmann, S. *Research on Direct Instruction: 25 Years beyond DISTAR.*; Educational Achievement Systems: Seattle, WA, USA, 1996.
- 72. Engelmann, S.; Carnine, D. *Theory of Instruction: Principles and Applications*; Irvington Publishers: New York, NY, USA, 1982.
- 73. Swanson, H.L. Searching for the best model for instructing students with learning disabilities. *Focus Except. Child.* **2001**, *34*, 1.
- 74. Braithwaite, J.J.; Watson, D.G.; Jones, R.; Rowe, M. A guide for analysing electrodermal activity (EDA) & skin conductance responses (SCRs) for psychological experiments. *Psychophysiology* **2013**, *49*, 1017–1034.
- 75. Benedek, M.; Kaernbach, C. A continuous measure of phasic electrodermal activity. *J. Neurosci. Methods* **2010**, 190, 80–91. [CrossRef]
- 76. Tarvainen, M.P.; Niskanen, J.P.; Lipponen, J.A.; Ranta-Aho, P.O.; Karjalainen, P.A. Kubios HRV-heart rate variability analysis software. *Comput. Methods Programs Biomed.* **2014**, *113*, 210–220. [CrossRef]
- 77. Baker, R.S.; D'Mello, S.K.; Rodrigo, M.M.T.; Graesser, A.C. Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive–affective states during interactions with three different computer-based learning environments. *Int. J. Hum. Comput. Stud.* **2010**, *68*, 223–241. [CrossRef]
- 78. Fowles, D.C. 10 The Measurement of Electrodermal Activity in Children. In *Developmental Psychophysiology: Theory, Systems, and Methods;* Cambridge University Press: Cambridge, UK, 2007; p. 286.

- 79. Bloom, B.S. Bloom's Taxonomy of Educational Objectives; Longman: Harlow, UK, 1965.
- Kharrufa, A.; Rix, S.; Osadchiy, T.; Preston, A.; Olivier, P. Group Spinner: Recognizing and visualizing learning in the classroom for reflection, communication, and planning. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, CO, USA, 6–11 May 2017; pp. 5556–5567.



 \odot 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).