Chapter 7

Measuring Electrodermal Activity in an Afterschool Maker Program to Detect Youth Engagement

Ryan Cain
Weber State University, USA

Victor R. Lee
https://orcid.org/0000-0001-6434-7589
Stanford University, USA

ABSTRACT

In this chapter, the authors describe a new approach for exploring individual participants' engagement in immersive youth maker activities. Participants were outfitted with wearable first-person point-of-view still-image cameras and wrist-based electrodermal sensors. The researchers analyzed the recorded electrodermal data stream for surges in skin conductivity and compared them with the corresponding photographs based on their timestamp. In following with prior work, these surges were interpreted as moments of engagement. A comparison sample was created to look at moments that lacked this psychophysiological marker. Results suggested that the two participants had both shared and divergent engagements with the afterschool program's activities. While the group project of building a high altitude balloon had been established prior to the youth's participation, the girls were able to choose what aspect of the project they wanted to be responsible for. This range of activities provided opportunities for youth to sample a variety of practices typically associated with making.

DOI: 10.4018/978-1-7998-3250-8.ch007

INTRODUCTION

Making is thought to be an immersive and highly engaging set of technological practices for a broad range of youth (Calabrese Barton, Tan, & Greenberg, 2016; Svihla, 2015). As such, there is an opportunity for researchers to better understand the ways that individuals differentially engage with making. This is important to explore because making often provides an opportunity for custom experiences. What one youth does in a makerspace may be different from what another youth does at the same makerspace. Moreover and relatedly, what one youth finds engaging in a makerspace may differ from what another youth finds engaging. This could apply when makers are doing the same activity. The current study aims to examine youths' psychophysiological responses to some common maker activities in the context of a multi-week afterschool maker education program. Typical maker activities include crafting with both traditional and digital tools to fabricate a designed object (Martin, 2015). In this study the participants collaborated on the building of a high attitude balloon and sensor payload. The current study investigated participants' psychophysiological responses while engaged with maker activities in the context of a project with a predetermined final product, the high-attitude balloon. This study examines whether there are indeed differential levels of engagement for youth experiencing the same activities in a makerspace.

In the proceeding sections we first provide a brief overview of prior work on the use of skin conductance as a psychophysiological measurement and follow with an explanation of the role that engagement and interest play in making. Next we present a new method we are exploring and initial findings from it using wearable cameras and sensors to examine engagement in maker activities at a minute-by-minute grain size. Finally, we conclude with a discussion on our interpretations of this new method, their potential for future work, and limitations.

BACKGROUND

Electrodermal Activity

The measurement of electrodermal activity (EDA), sometimes referred to as skin conductance or galvanic skin response, gauges psychophysiological activity of the sympathetic nervous system. That is, we are attending to the response of a physiological body system to psychological states and changes. Orienting reflex theory explains the activation of the sympathetic nervous system as orienting response, where "OR (orienting response) functions to produce a heightened sensitivity to environmental stimulation and results in increased intake and processing of information" (Raskin, 1973, p. 128). This response prepares the body for action by signaling sweat glands on or near the hands and feet to produce sweat to increase friction for a better grip and traction respectively (Matsumoto, Walker, Walker, & Hughes, 1990).

Electrodermal activation may not produce a visible amount of sweat at the surface of the skin. However, the electrical conductivity of the skin increases as the sweat glands begin producing sweat into the sweat ducts as part of the autonomic response (Boucsein et al., 2012). The skin can be thought of as a sponge that becomes increasingly conductive as salty water fills the ducts that connect the sweat gland to the surface of the skin. This change in skin conductivity can be measured by passing a small current through the skin and measuring the resistance. The value is presented as the inverse of the resistance with the unit microSiemens (μS).

Sensing EDA has long been a tool for measuring human affective psychophysiological responses to stimuli (Matsumoto et al., 1990). Pecchinenda (1996) measured surges of EDA to indicate engagement while participants were attempting to solve Tangram puzzles. The participants worked on solving puzzles of varied difficulty levels and the researchers found that participants stopped engaging as the puzzles presented became increasingly difficult. This corresponded with a decrease in EDA. Kreibig, Gendolla, and Scherer (2012) similarly found a rise in EDA when their participants were given relevant feedback to help with the achievement of a goal. Both of these studies used EDA recording devices that required participants to be attached to a desktop computer via wires instead of moving about freely.

The current study relies on wearable EDA recording devices to collect data unobtrusively allowing wearers to have full use of their hands without electrodes on their fingertips (Poh, Swenson, & Picard, 2010). Traditionally, EDA is measured with more obtrusive skin sensors with electrodes tethering two fingertips to a computer with wires. It is important to note that finger-based sensor would interfere with the types of activities conducted in a makerspace requiring full use of the hands. Furthermore, participants needed to be able to move about the makerspace in order to access various tools and components.

The current work differs further from prior work with how the task being observed was neither predetermined nor did it all occur on a computer screen. Previous work examining affective responses of learners involved participants working on computers that captured both physiological measures and the learning activity being conducted on the screen (Arroyo et al., 2009; Blikstein, Gomes, Akiba, & Schneider, 2016), which were predetermined activities. The current work had participants wear first-person point-of-view still-image cameras to document what the student was seeing and hearing, documenting the unplanned and mobile work. At times they might be using a hand tool, listening to a mentor explain how to fix a non-functional circuit, or sending a text on their phone. These paired data streams provide a novel perspective into wide range of things people do while making by simultaneously monitoring external actions and internal psychophysiological activity (see Lee, Fischback, & Cain, 2019 for another example of this methodological approach).

Making AS an Engaging Activity

Ongoing engagement with a maker project may lead to desirable outcomes. Fields and King (2014) share one such example where an undergraduate student in a university-level maker course finds success with the unfamiliar content of programming to the point where she develops an interest and an emergent identity as a programmer. Brahms and Crowley (2016) provide another example of repeated engagement in making with a four-year-old boy working in a museum-based makerspace over multiple visits. While initially unsure about what to make, he built a series of toy power tools under the support of museum educators and his mother. The museum educators elicited the boy's preexisting interest in his grandfather's power tools and used it as a source for inspiration of what to make. The museum educators skillfully leveraged the child's established interest in order to motivate the child to be engaged with the maker activities in the museum. This example demonstrates the utility of enlisting a preexisting interest to promote engagement in an activity. Since we cannot always count on a one-to-one mentor being available or able to solicit established interests and harness them to promote the development of new interests in maker education contexts, additional tools are needed to inform the support of young makers. Cain, Phillips, & Lee (2018) document a case of a youth's engagement with a laser cutter over an extended period of time as an interest develops from initial contact with the laser cutter to eventually teaching other youth how to use the laser cutter. In all these examples, interest develops from where there had seemingly not been an interest in making previously. The current work seeks to interpret psychophysiological responses for the potential detection of nascent interests as they are forming.

Engagement as a Phenomenon of Interest

Engagement as a psychological and interactional phenomenon relies upon intuitive appeal. We are all personally familiar with the cognitive and affective state of high engagement from our own experiences. From those, we believe engagement typically involves high levels of attention, intention, participation, and increased affect (both positive and negative). Yet how engagement is operationalized and recognized is still under-specified and subject to overarching theoretical perspectives that suggest how engagement is framed in relation to a given task or setting (Sinatra, Heddy, & Lombardi 2015).

For the present study, we have drawn on a more established, but still developing body of work related to interest to situate engagement. One of the most recognized theoretical models of interest comes from Hidi and Renninger (2006). Their four-phase model of interest development posits that interest progresses in the following stages: phase 1—triggered situational interest; phase 2—maintained situational interest; phase 3— emerging individual interest; and phase 4—well-developed individual interest. To illustrate, we could imagine a youth enjoying watching a dinosaur program on her mobile device that automatically loaded after her favorite show about child veterinarian—phase 1. Upon a visit to the museum the child asks to visit the dinosaur exhibit first where she spends several awestruck minutes at the exhibit—phase 2. This interest develops further such that while at the local public library, this youth checks out several dinosaur books and asks the librarian where to find additional resources—phase 3. For career day at her school she presents a project about her future self as a paleontologist specializing in researching dinosaurs and is known among her peers to be someone who is seriously interested in dinosaurs—phase 4. In this example, the salient feature distinguishing phases 1 and 2 from phases 3 and four is that situational interests are externally triggered, while a person with an individual interest will seek out situations to engage the interest. For this chapter, we are most concerned with the externally triggered early stage of interest development, situational interest.

Drawing on the four-phase model, moments of engagement would align with moments of situational interest. That is, features of an activity that capture attention and elicit some seconds of consideration and immediate participation from a youth are times we expect to see momentary engagement. It is entirely possible, and we suspect, expected that over many repeated encounters, those moments of engagement may become more enduring. Once an interest has sufficiently developed such that we ascribe a particular interest to a person, then we would expect to see that person's continued and recurrent participation over longer periods of time with gaps in between. It is at that point we might say "Joyce is really interested in dinosaurs." In extended and unbroken periods of engagement, attentional focus would be such that a person would lose track of time while being fully absorbed in the task. In the example of Joyce and dinosaurs, that might come about during several minutes of reading about the Apatosaurus genus or sketching a T-Rex.

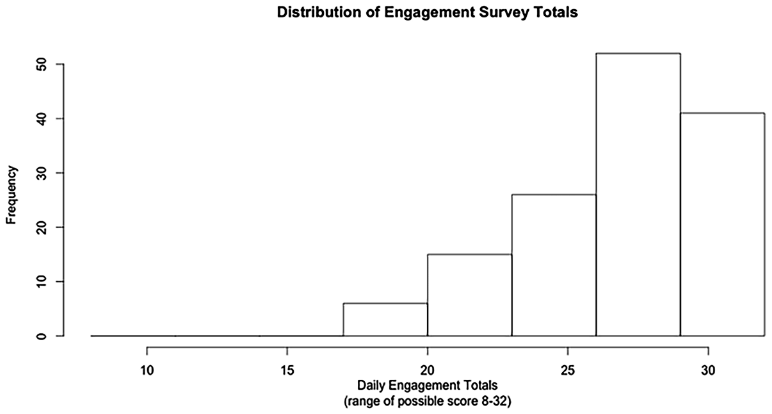
However, it is the earliest two phases of interest development that involve situational interest. Situational interest is something we posit will produce moments of heightened psychophysiological arousal. That arousal produces markers in the form of relatively immediate changes in skin conductivity. While there remains much still to be understood about skin conductivity changes in situ, this study explores what can be revealed when rapid increases in skin conductivity, a signature of physiological arousal, are used as a proxy and marker for moments of situational interest and engagement. While we present our method inferring engagement from EDA, we must acknowledge that this relationship is tentative. We do not know with certainty that engagement is associated with electrodermal activity changes. We simply rely on the same assumptions made by prior work (e.g., Daily et al., 2015; Pecchinenda, 1996). This paper presents our design for a new methodology based on the assumption of this proxy relationship between electrodermal activity and engagement.

Measuring Engagement

Although not a primary data source in the current study, as we are focusing on skin conductivity, we do wish to mention that we administered a more traditional instrument, an engagement survey, to the participants at the conclusion of each session to capture their overall level of engagement. This provided a more coarse picture of engagement in the makerspace, with individual days in the makerspace being the unit of analysis. The survey included 8 Likert items to measure students' affective-behavioral engagement and perceived success in each session's activities. This resulted in 145 daily surveys (12 maker sessions and 1 reviewing balloon data). A four-point scale was given for each item consisting of YES!, yes, no, and NO!. This instrument had been psychometrically validated in science classrooms and museum visits (Ben-Eliyahu et al., 2018). We observed a ceiling effect with the participants consistently reporting being highly engaged, M=27.44, SD=3.18 with 8-32 as the possible range of scores, see image 1 below. While this was encouraging to see high engagement self-reported at the end of every session, it does not tell us what specific activities in a session were more or less engaging. The method we use focuses on a smaller grain size of engagement, on the scale of seconds and minutes rather than an entire session. Our focus was on what moments during the session stood out according to our interpretation of individual students' electrodermal responses. However, the survey results suggest this population was highly engaged and thus there should be some detectable moments of engagement in the larger data corpus.

To get at moments of engagement, the present study utilized wearables (skin conductivity sensors) to measure arousal of the sympathetic nervous system. First person perspective images allowed for the identification of what the participants were behaviorally engaged in, and skin conductivity was interpreted as a proxy for when youth were situationally engaged. Together these primary data sources were analyzed to examine what participants were doing while their skin was exhibiting a psychophysiological

Figure 1. Total Engagement score distribution of 145 surveys



response. We chose this over other methods such as experience sampling, where a participant might receive a notification on a mobile device asking about their engagement in an activity (Schneider et al., 2016), as not to interrupt the engagement we sought to understand. Another approach would have been to pursue multimodal learning analytics which utilizes various unobtrusive methods of data collection, including eye-tracking, skin conductivity, facial expression, hand gestures, and posture, to study learning in messy environments (Andrade & Danish, 2016; Worsley, Scherer, Morency, & Blikstein, 2015). Those techniques are still under active development to better understand these messy environments in more automated ways, as opposed to labor intensive direct observation of individual learners. The current work is more modest and seeks to tease out the potential utility of skin conductivity – one stream of data that is used in multimodal learning analytics – as a means for studying the cognitive and affective phenomenon of engagement.

METHOD

Setting and Participants

Twelve female participants attended a twelve-week space-themed after-school maker club at a youth makerspace in a rural area of the intermountain west in the United States. The current study focuses on two of the participants,[1] ages 10 and 12 whom we will refer to as Dot and Jane[2]. These two girls worked together as partners. They participated in the same activities side-by-side for the vast majority of their time in the makerspace.

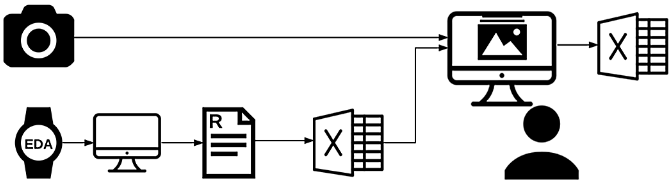
The maker club was structured so that all 12 girls met once a week for two hours over the course of 12 weeks. Parent volunteers and makerspace employees mentored the girls with a child to adult ratio of roughly 4:1. (For a history of and information about the structure of the makerspace, refer to Lee, King, & Cain, 2015.) During these meetings they worked in pairs building and testing an array of sensors to be launched into the upper atmosphere. In addition to the wearable data described in the following section, we collected digital video from a tripod-mounted camera directed at Dot and Jane and daily engagement surveys administered to all 12 youth members of the afterschool club, described above.

The group followed a combination of activities developed by the group's leader, Christopher, and the Ardusat online curriculum (ardusat.com/lessons, a space themed experiment platform for facilitating physical-computing based inquiry). These activities included building antennas for locating hidden radio transmitters, wiring sensors, programming Arduino microcontrollers to read sensor values, calibrating sensors, and constructing an enclosure to transport the electronics under the high-altitude balloon. Each participant pair was responsible for an individual sensor having to wire, test, and interpret the resultant data. The sensors measured luminosity, ultraviolet light, infrared light, acceleration, orientation, and temperature. During the final weeks they installed the sensors, a radio beacon transmitter, a radio GPS transmitter, and an action camera into three 7" x 7" x 7" foam boxes. The young women launched and recovered their sensor-laden payload after it traveled to an altitude of over 110,000 feet and a distance that traversed a mountain range and state line.

Data from Wearable Devices

For this chapter we present data from two partnered participants over three meetings of the after-school club. The three sessions were selected for their breadth of maker activities, including the use of hand tools, programming and wiring of Arduinos, and testing sensors. We chose two out of the four girls who wore wearables since they worked as a pair and were in attendance at all the sessions. The goal of the following method was to establish a comparison between instances of rapid surges of skin conductivity and a comparison sample drawn from periods of relatively constant levels of skin conductivity. A spike or sudden rise in skin conductivity indicates arousal of the sympathetic nervous system, as mentioned earlier. We use the terms arousing moments and unarousing moments to categorize times of rising and relatively flat levels of skin conductivity respectively. In the following sections we explain our procedures for preparing our data for analysis, first by detecting the above features and then pairing the temporally matched images to the moments. The time scale of a moment ranges from less than a minute to several minutes depending on the duration of the rising EDA. An overview of the workflow is depicted in figure 2.

Figure 2. Electrodermal activity data (bottom) is processed to create two sets of data for comparison: one with rapid surges in skin conductivity (arousing moments) and one with periods of constant skin conductivity (unarousing moments). The resultant data sets are paired with their temporally matched images from the wearable cameras (top). Those images are then coded for arousal state and content (objects, setting, people, and actions).



Finally the codes are summarized in a spreadsheet for the comparison analysis.

Table 1. List of primary and secondary data sources

	Data Source	Quantity
Primary	Wearable skin conductance sensor	Approximately 57,600 EDA values
Primary	Wearable point-of-view camera	Approximately 4,500 images
Secondary	Tripod mounted video camera	Approximately 6 hours
Secondary	Engagement survey given at the conclusion of each session	145 responses from all 12 youth in makerspace

Our primary sources of data were wrist-based EDA sensors, *Affectiva Q*, and point-of-view still-image cameras, *Autographer*, worn around the participants' necks, see Figure 3 below. Skin conductivity (EDA) was recorded at 4Hz resulting in about 57,600 values total for the two youth over the three days. The cameras captured an image once every 8-15 seconds based an internal algorithm interpreting motion and lighting conditions yielding about 4,500 images. The wearable devices allowed the participants to freely move about while capturing time-stamped EDA and still images.

Figure 3. Autographer camera and two Affectiva Q sensors.

EDA Data Processing

In this section we describe our process for identifying surges in skin conductivity and establishing a comparison set of intervals of relatively flat skin conductivity, resulting in sets of arousing moments and unarousing moments. Each of the two participants' wearable EDA sensors generated approximately 24,000 values per session—4Hz for 2 hours (see figure 4 below). We applied the following algorithm in R to develop a consistent way for identifying spikes—indicating EDA arousal—while reducing noise. The arithmetic mean was calculated for the 40 EDA datum collected per 10-second interval—chosen to approximate the wearable cameras interval of a photo every 8-15 seconds—resulting in about 600 intervals per day. Each interval was subtracted from the following one in a $\Delta = X_{i+1} - X_i$ calculation resulting in what we will refer to as relative change in skin conductivity (RCSC). In simpler terms, the RCSC is the slope of EDA plot after reducing the data from a value every 0.25 seconds to one every 10 seconds. The 599 RCSC values are shown as dots on the bottom of figure 4.

In order to detect spikes in skin conductivity, we chose one positive SD of the distribution of RCSC values as the threshold for defining arousing moments, indicating rising EDA to capture positive RCSC values not near zero. A randomly selected comparison set of unarousing moments was selected from RCSC values that were within ±0.25 SD because this selected values near zero, see bottom of Figure 4. The resulting analysis contains two categories of moments, arousing and unarousing, shaded in light gray for arousing and dark gray for unarousing on the bottom of Figure 4.

Since our data reduction method defined arousing as greater than one SD and unaroused as within 0.25 SD of the mean, we have excluded all other RCSCs that do not fit the criteria, seen in figure 4 as dots not shaded in light or dark gray. This was a decision to conservatively define arousal as the largest EDA surges excluding rises less than 1 SD. This threshold for spike detection identified the largest spikes (the steepest positive slopes) which we later infer as the possible initiation of an episode of situational engagement. Additionally, the intervals of descending EDA greater than -0.25 SD are excluded as returning to baseline period. The quantity of unarousing moments comparison sample was set to be equal to the number of arousing moments for the participant with the greater number for that day. We based the comparison sample on the larger quantity as to reduce bias that could be introduced with the smaller random sample between participants. For example, on 8/11 Jane had 61 arousing moments and Dot had 51, the comparison unarousing moment set would be from 61 random moments ±0.25 SD for both participants. These unequal comparisons within participants are accounted for in the image coding section below.

Figure 4. (Top) Electrodermal Activity for Dot on 9/8 (Bottom) relative changes in skin conductivity (RCSC) between 10-second intervals for Dot on 9/8. Light gray shading indicates arousing moments; dark gray shading indicates unarousing moments according criteria specified in the current study.

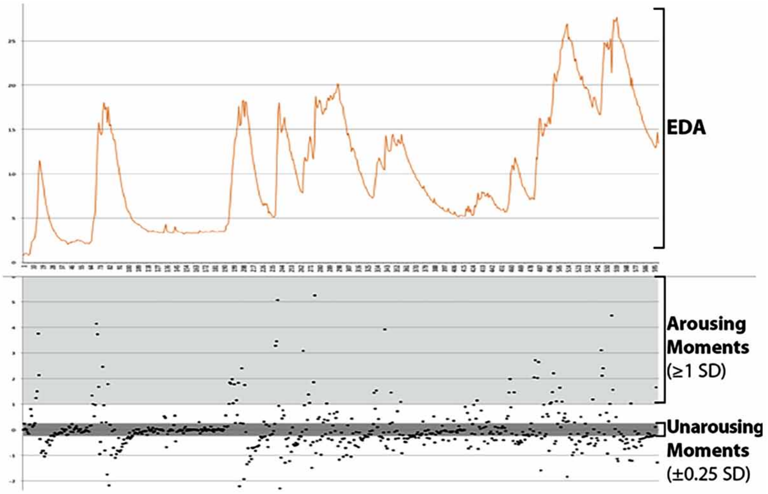


Photo Inclusion Criteria

Having established sets of arousing and unarousing moments (i.e., high and low engagement moments), we then used the time-stamps associated with RCSCs and first-person photographs to match images with moments. Any image taken during a minute that contained an arousing or unarousing moment was copied to a photo database for coding. The number of photos taken per minute varied due to the camera's decision-making algorithm, but at a minimum of one photograph every 15 seconds. The resultant image sets, aroused and unaroused images, became the basis of our comparison of periods of rising skin conductivity to relatively flat ones.

Table 2. Partial list of coding scheme for image analysis

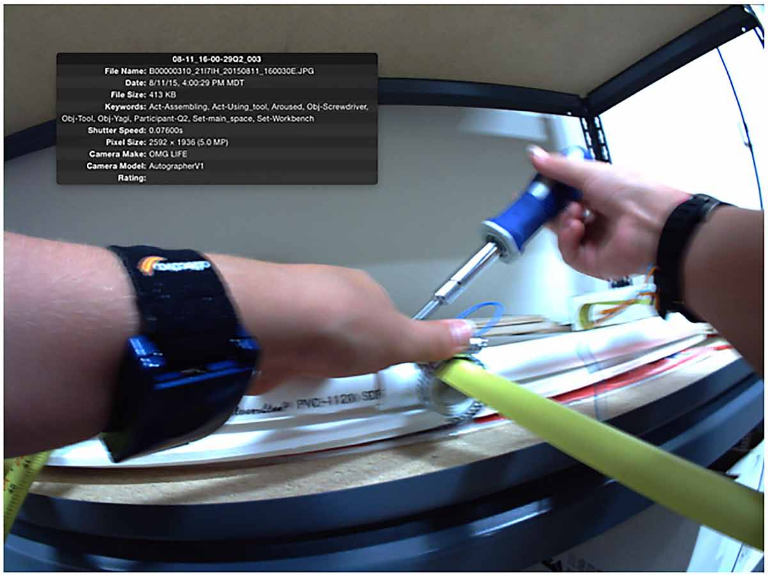
Objects	Setting	People	Action	Arousal State
Arduino	Main makerspace	Participant 1	Assembling	Aroused
Drill	Lobby	Participant 2	Testing	Unaroused
LED	Outside	Dot	Eating treat	
Laptop	Workbench	Jane	Programming	
Saw	Woodshop	Mentor 1	Hands resting	
Screwdriver	Supply rack	Mentor 2	Listening to lecture	
Soldering iron		Dot's Mother	Watching someone model	

DATA ANALYSIS

Coding Photos

We developed a coding scheme to analyze the first-person perspective images from the wearable cameras in order to interpret what was happening in the photographs; see table 2 for a sample of those codes. Our process for creating this scheme involved the authors viewing contact sheets (thumbnails of all the images printed in color as a 5 x 6 arrays on 11" x 8.5" paper) identifying the materials used by the participants in the makerspace, noting the different work areas in and surrounding the makerspace, and discussing who the youth interacted with and what they appeared to be doing. Several of these shared viewing sessions allowed us to refine our coding scheme. One such example for image 5 below follows:

Figure 5. Coded image of Jane assembling an antenna with a screwdriver. The EDA wearable can be seen on her left wrist.



08-11_16-00-29Q2_003
File Name: B00000310_21I7IH_20150811_160030E.JPG
Date: 8/11/15, 4:00:29 PM MDT
File Size: 413 KB
Keywords: Act-Assembling, Act-Using_tool, Aroused, Obj-Screwdriver, Obj-Tool, Obj-Yagi, Participant-Q2, Set-main_space, Set-Workbench
Shutter Speed: 0.07600s
Pixel Size: 2592 × 1936 (5.0 MP)
Camera Make: OMG LIFE
Camera Model: AutographerV1
Rating:

originally, we coded it as *object-screwdriver* and *action-using tool* but later refined the coding to also include *action-assembling*. Having gone through the first pass of coding made the categories of actions taking place more visible.

The coding identified objects in the frame (Arduino, tool, laptop), setting (outside, workbench, walking), people (mentor and participant names), and the actions taking place (*assembling, testing, programming*) when identifiable from the images. The action codes served as a summary of the moment when identifiable and as our object of analysis in later sections. These codes were included in the image files' metadata and subsequently analyzed in a spreadsheet to compare quantities across participants and days.

Comparing Code Counts

We decided to compare code counts between periods of relatively flat EDA to times of surging EDA to determine if codes were distributed evenly between these two conditions, or if certain codes were more associated with a certain state. We chose a threshold proportion of at least three aroused moments for every two unaroused moments to categorize a particular activity code as a typically arousing activity. This was to build in a criterion for determining if an activity was arousing, instead of simply stating there were more aroused moments than unaroused. The intervals were not directly coded; rather photographs from the wearable cameras taken during the minute of the RCSC were coded. This is why 61 aroused RCSCs for Jane on August 11 resulted in 264 coded images. We established a comparison sample of unaroused moments equal to 61. In order to make comparisons between the participants, we based the unaroused sample on the participant with the greatest number of spikes; in the above example this was 61 unaroused RCSCs for both Jane and Dot, even though Dot had 51 aroused RCSCs. Once again, this was an attempt to build in a buffer by sampling the larger portion of the unaroused moments. Having these unequal samples due to our method and the wearable camera's algorithm for picture frequency, we utilized proportional scaling to make fair comparisons explained in the following paragraph.

Comparing code counts served as a way to infer what activities were more or less engaging as measured by spikes of skin conductivity. As a reminder, we defined a spike in skin conductivity as relative change in skin conductivity (RCSC) from one 10-second interval to the next 10-second interval greater than or equal to +1SD of all the RCSCs for that participant on the specific day. For example, this criterion identified 61 and 51 spikes for Jane and Dot on August 8. In order to make comparisons between spikes and periods of relatively flat RCSC, we established a comparison random sample drawn from all of the unaroused moments as defined as RCSCs within ±0.25 SD of the minimum for that child on the day. The number of randomly selected unaroused moments was selected so that there were equal quantities of aroused and unaroused images for each participant each day.

First, we created code count totals by day for the arousing and unarousing images separately. Next we applied proportional scaling to the unaroused counts to adjust for the unequal comparison within participants. This unequal comparison was caused by variation in the number of pictures taken per minute and by our choice to create the unarousing moment sample equal to the number of arousing moments of participant with greater quantity. In table 3, Jane's 61 arousing RCSCs on August 11 served as the reference for selecting both Jane and Dot's quantity of unarousing RCSCs for that day. Looking across the image totals on August 11, we point out that using 61 RCSCs as the reference for the date (see 8/11 row in table 3) the photo inclusion criteria identified 264 and 251 images for Jane (aroused and unaroused respectively) and 245 and 217 images for Dot. The proportional scaling allowed for equal comparisons within participants' arousing and unarousing image counts by individual day. Next we turned these

Table 3. Quantities of arousing and unarousing RCSCs with quantities of corresponding of images in parentheses

	Jane		Dot	
	Arousing	**Unarousing**	**Arousing**	**Unarousing**
8/11	61(264)	61(251)	51(245)	61(217)
9/8	26(54)	48(210)	48(155)	48(214)
9/15	40(127)	40(184)	37(150)	40(179)

comparisons into ratios to compare between days and participants. The proportional scaling within a ratio of unaroused to unaroused images was calculated with the following equation:

Note, the scaling factor is bold. Having established equal comparisons within participants per day, we applied this scaling factor across all counts of unaroused code totals and calculated the **arousal ratio** (AR) for code by participant/day. AR, the ratio of unaroused to aroused images, enabled comparison between and within participants by day.

Next we defined a threshold AR value to categorize what activities were proportionally arousing by comparing counts of arousing and unarousing images for each activity. Events with a ratio three arousing moments to two unarousing moments or less ((unaroused event totals)/(aroused event totals) ≤ 0.66) were deemed **proportionally arousing**, and AR > 0.66 were **proportionally unarousing**.[3] For example, we coded for the activity *Assembling* with 47 aroused images and 16.94 unaroused images (this decimal count resulted from the scaling factor) for Dot on August 11 that yields an arousal ratio of 0.36, meeting our criteria for a proportionally arousing activity. We used these ratios to look across activities over the three days at where the two participants' arousal ratios (AR) coincided and diverged.

RESULTS

Having calculated arousal ratios for the various coded activities for both participants over the three days, we began to look for the presence of similarities and differences between the girls. While we used conditional formatting in Excel for the analysis, the heat map below makes for a clearer overview of the ratios with crosshatched red indicating proportionally arousing activity and box size denoting the number of images the ratio is based on. In the following sections we first share activities where the dyad exhibited contrasting ARs (as seen when their colors don't match below in figure 6) to activities and follow with activities where they had similar reactions (matching colors). We conclude with a section on activities where the pair had similar responses with a secondary analysis to better understand moments of arousal during the proportionally unarousing activities of programming and testing.

Contrasting Psychophysiological Responses

As expected, we observed times where the Dot and Jane's ARs differed. On 8/11 the girls spent part of the session constructing an antenna, during which they *watched someone model* how to use a soldering iron, *assembled* parts, and *soldered* wires. We noticed that the two young women exhibited opposite ARs on these four activities with solid blue and crosshatched red boxes on the heat map in figure 6.

Figure 6. Heat map depicting relationship between images coded as aroused and unaroused. Cross-hatched red indicates proportionally aroused and solid blue indicates proportionally unaroused. Box size represents the quantity of images coded for that activity.

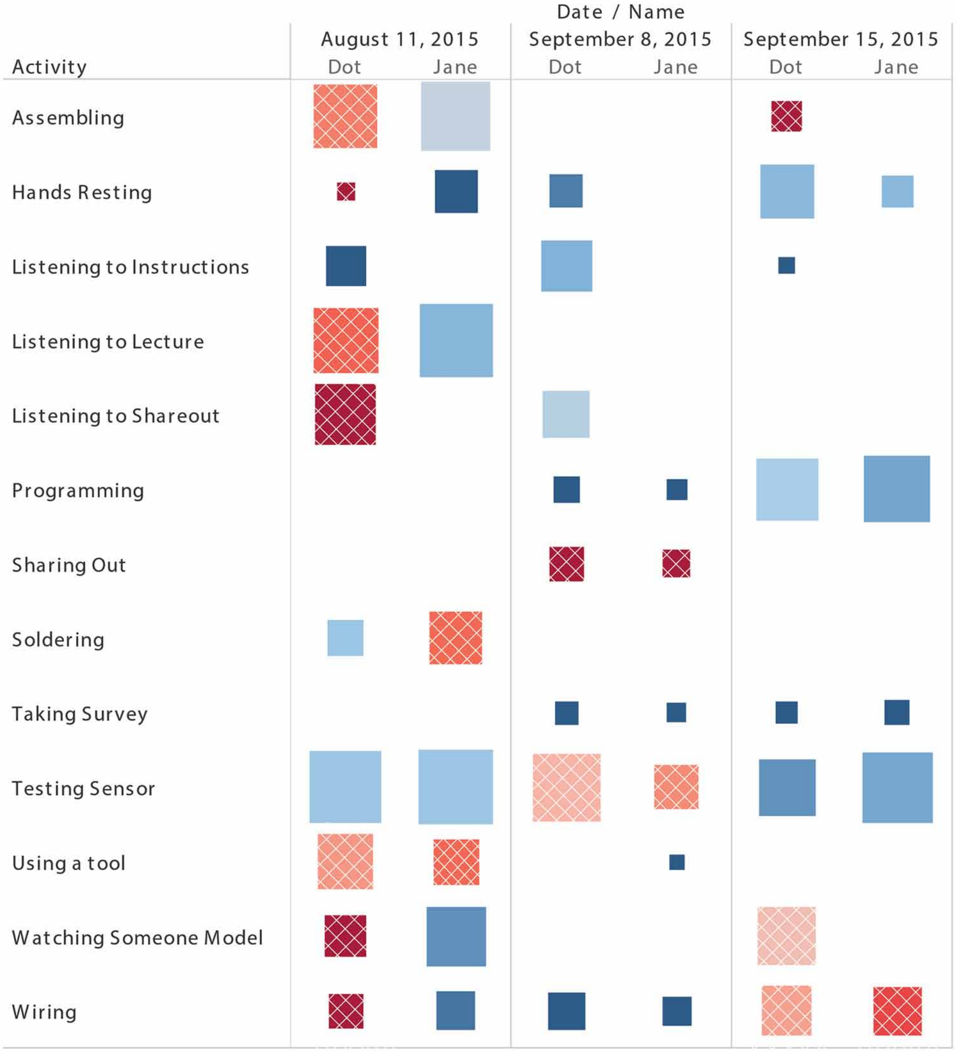
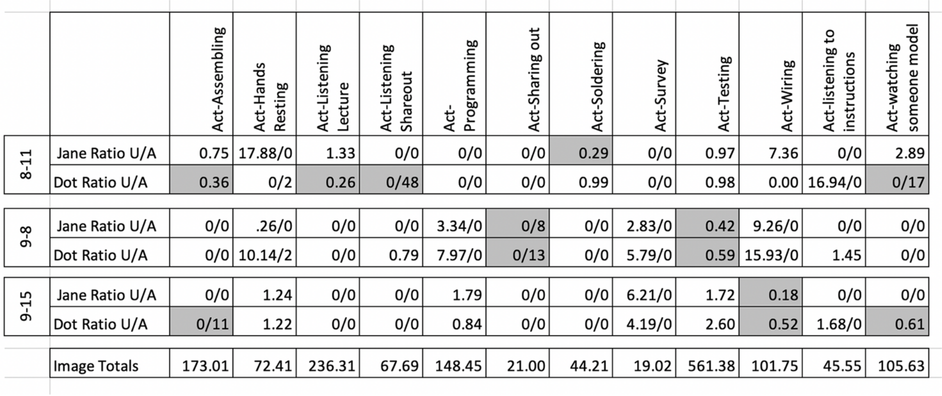


Figure 7. Partial summary of arousal ratios for Jane and Dot over 3 days. Proportionally arousing activities are shaded in gray. A limitation of our arousal ratio surfaces when a zero value appears in the ratio. In these instances, we interpret the fraction instead of the decimal form as being exclusively arousing or unarousing for that day.



		Act-Assembling	Act-Hands Resting	Act-Listening Lecture	Act-Listening Shareout	Act-Programming	Act-Sharing out	Act-Soldering	Act-Survey	Act-Testing	Act-Wiring	Act-listening to instructions	Act-watching someone model
8-11	Jane Ratio U/A	0.75	17.88/0	1.33	0/0	0/0	0/0	0.29	0/0	0.97	7.36	0/0	2.89
	Dot Ratio U/A	0.36	0/2	0.26	0/48	0/0	0/0	0.99	0/0	0.98	0.00	16.94/0	0/17
9-8	Jane Ratio U/A	0/0	.26/0	0/0	0/0	3.34/0	0/8	0/0	2.83/0	0.42	9.26/0	0/0	0/0
	Dot Ratio U/A	0/0	10.14/2	0/0	0.79	7.97/0	0/13	0/0	5.79/0	0.59	15.93/0	1.45	0/0
9-15	Jane Ratio U/A	0/0	1.24	0/0	0/0	1.79	0/0	0/0	6.21/0	1.72	0.18	0/0	0/0
	Dot Ratio U/A	0/11	1.22	0/0	0/0	0.84	0/0	0/0	4.19/0	2.60	0.52	1.68/0	0.61
	Image Totals	173.01	72.41	236.31	67.69	148.45	21.00	44.21	19.02	561.38	101.75	45.55	105.63

Looking at Figure 7, Jane and Dot differ on August 11 with arousal ratios (AR) of 0.75 and 0.36 respectively for the activity *assembling*. Dot was *assembling* a sensor on September 15 where she had 11 arousing moments (0.00 AR). While Dot was proportionally aroused by the act of *assembling* both on August 11 and September 15, she was less aroused by the activity of *soldering*. In contrast Jane was proportionally aroused while *soldering* (0.29 AR), but not quite during *assembling* (0.75 AR). Dot was split evenly between arousing and unarousing moments while coded for *soldering* (0.99 AR). Stated another way, Dot appeared to be more engaged with *assembling* while Jane was more engaged with *soldering*.

During the activity of *listening to lecture* on August 11, we see Dot was proportionally aroused with a 0.26 AR, while Jane was proportionally unaroused with 1.33 AR. On the same day, the code *watching someone model* shows a similar arrangement with Dot with 0.00 AR (0 unarousing moments/11 arousing moments). Additionally, we point out that Dot on a separate day, September 15, has images coded for *watching someone model,* as in modeling how to solder, were proportionally arousing with a 0.61 AR. Our analysis of these four activity codes (*assembling, soldering, listening to lecture, and watching someone model*) begins to reveal some differences between how the pair of girls reacted differently to the shared activities. That is, despite participating in the same activities together, they exhibited differing degrees of engagement in comparison to one another.

Table 4. Arousal ratios (AR) for Jane and Dot while programming and testing sensors.

	Programming		Testing Sensor	
	Jane	**Dot**	**Jane**	**Dot**
8/11	0/0	0/0	0.97	0.98
9/8	3.34/0*	7.97/0*	0.42	0.59
9/15	1.79	0.84	1.72	2.60

Note. *Denotes AR ratio indicating proportionally arousing activity

Shared Psychophysiological Responses

There were times Jane and Dot shared a common response. For instance, this happened on September 8 with ARs of 0/8 and 0/11[4], for images coded as the *activity sharing out*. During those moments the pair stood in front of their peers and mentors explaining their progress, so it is not surprising to see rises in EDA when it is common for presenters' hands to sweat. Here we are pretty confident that they were engaged while sharing out, but whether that was out of fear from presenting, being excited about talking to their peers about what they worked on that day, or some other factors remains unknown. Other common responses for the pair included the *activities-programming* and *testing sensor*. Since these two activities are central to the practice of making and for the most part were proportionally unarousing, we chose to make a second analysis to sift out other factors that may explain times of arousal within these activities.

The participants in this study mainly utilized preexisting computer programs from the online Ardusat curriculum (ardusat.com), only making minor modifications such as changing the input that a sensor was connected to. We chose September 15 since Jane and Dot both were programming on this day with corresponding ARs of 1.79 and 0.83 based on a total of 137.14 coded images. While we acknowledge that their ARs are not equal, they are both above our 0.66 proportionally arousing moment threshold indicating that programming was proportionally unarousing. Similarly, on September 8 programming was unarousing for both girls, but this was based on a small number of exclusively unarousing images coded for both girls. After reviewing the coded images for *programming* on printed contact sheets, we noticed images of arousing moments while *programming* often had a mentor in the image. We confirmed this observation with the code counts, see table 5.

Table 5. Distribution of image counts for the activity Programming by the presence a mentor.

	Mentor Present		**No Mentor**	
	Images	**Percent**	**Images**	**Percent**
Arousing	34	40.48%	25	23.36%
Unarousing	50	59.52%	82	76.64%
Total	84	100.00%	107	100.00%

We examined video as a secondary data source to better understand what role the mentors were playing during programming, which showed there were two particular makerspace-employed mentors that

Table 6. Image counts breaking down the activity of testing sensor for Jane and Dot combined

	Inside		**Outside**	
	Images	**Percent**	**Images**	**Percent**
Arousing	85	40.21%	191	54.57%
Unarousing	126.4	59.79%	158.99	45.43%
Total	211.4	100.00%	349.99	100.00%

using being indoors or outside the makerspace as a factor

Figure 8. Mentor working with Jane to alter how a program displays a sensor value

were helping Jane and Dot to alter code provided by tutorials. In one instance Jane wanted to modify the behavior of the preexisting program that read and displayed the value from an ultraviolet light sensor. Although the program worked, Jane wanted the display of the value to not keep scrolling up the screen. We posit that the mentors' one-on-one help while programming (see Figure 8) could evoke a surge in skin conductivity as a possible explanation for a greater percentage of the images being coded as being arousing and with a mentor 40.48% as opposed to only 23.36% for images coded as being arousing and no mentor present. We infer that working with mentors on an individual basis was engaging for these youth.

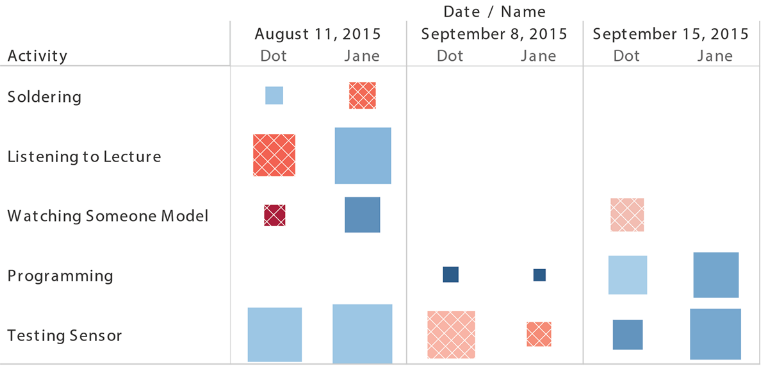
Next we took a similar approach to examine the activities coded as *testing sensor*. Images coded for *testing sensor* outnumbered the next most numerous code of *assembling* by more than 3:1. About 28% of all the coded images (561.38/1990) for both participants have been coded as testing, which was not surprising given that making involves repeated testing until the made object works properly. *Testing sensor* spanned across all three days with a shared response for Jane and Dot with August 11 and September 15 being proportionally unarousing and September 8 being proportionally arousing as seen in table 4. Looking for another factor in our image code counts, we recognized that the distributions of aroused and unaroused moments differed when Jane and Dot were outside as compared to being indoors, see table 6.

This could be an indication that getting out into the field to test a sensor is more arousing than staying inside the makerspace. While there is a significant difference of the distribution of inside *testing* images (X^2 *(1, N=211.40)* = 7.967, *p*=.005) and no significance outside (X^2 *(1, N=349.99)* = 2.926, *p*=.087), we acknowledge that it is quite possible that being outside may be generally more arousing without necessarily making testing outside more arousing. It is possible the less controlled climate conditions led to greater electrodermal activity, although we do not rule out that testing sensors under more authentic conditions may have been more arousing in general.

DISCUSSION

Based on these three days of data, we notice distinct profiles emerging from Jane and Dot's arousal ratios during maker activities. For Jane we pointed out earlier that *soldering* stood out as an engaging activity. Unlike Jane, Dot was more engaged during a lecture and while watching a mentor model how to complete a task such as using a soldering iron; see Figure 9 below for a summary of ARs discussed here. This suggests different potential pathways for interest development for these two girls. While still speculative, it may be that these data show that one girl (Jane) may be more likely to develop interest through hands-on tool use because those activities contained more moments of situational engagement and thus triggered more occasions of situational interest. The other girl (Dot) may be more inclined to develop a new interest through the presentation of information from a mentor figure, as there were more moments of situational engagement for her during those kinds of experiences.

Figure 9. Excerpt of heat map depicting arousal ratios (ARs). Crosshatched red indicates proportionally aroused and solid blue indicates proportionally unaroused. Box size represents the quantity of images coded for that activity.



Despite the observation that the girls both exhibited proportionally unarousing ARs for programming and two days of testing sensor activities, we observed that setting (indoors vs. outside) and the presence of a mentor might have an effect on arousal, which makes sense intuitively to also apply to a person being more engaged when being outside exposed to less static environment and interaction with a mentor on a one-to-one basis respectively.

Returning back to our goal of measuring engagement, we argue that exhibiting a surge in skin conductivity while participating in a maker activity may be indicative of momentary engagement, a way of engaging that accommodates the body's orienting response calling for a heightened awareness influencing the cognitive and affective components of engagement. We also caution that this would not capture all moments that a maker is engaged, rather that it is worth noting that instances of arousal should be investigated further. This work is important for both research on making and practitioners of maker education. In terms of research, we have proposed a method to be used to further test the utility of wearable-based electrodermal monitoring to begin to operationalize engagement and interest development in the messy real-world context of makerspaces. With further research this could be developed into a real-time formative feedback to help practitioners guide their learners' interest development within the scope of maker activities. Similar tools have been proposed for using electrodermal activity to inform teachers about states of engagement (Daily, James, Roy, & Darnell, 2015). Using Hidi and Renninger's (2006) model, during the first phase of interest development the learner "may or may not be reflectively aware of the experience" and "may experience positive or negative feelings" (Renninger & Hidi, 2016, p.13). In such a case where a person does not realize they are beginning to develop an interest, it would be advantageous to detect it using new tools so that appropriate activities can be provided so that interest has more opportunities to sustain itself. Future wearable devices could potentially detect a user's skin conductivity and activity in order to infer situational interest. Therefore, the detected situational interests could be used to provide users with suggestions for follow-up activities, possibly further developing a nacent interest.

We also note that the information about engagement gleaned from this method should be used with a specific directionality in mind. Our main goal was to test new technologies to gain insight into maker activities and how effective they are at providing experiences that engage learners. This ultimately provides feedback to designers of maker learning activities for things they may wish to do differently in a makerspace to accommodate and respond to different youth and help support the development of new interests. We do not advocate using these technologies to assess learners. For instance, we would consider it problematic if individual youth were critiqued or judged for failing to show a physiological response to a given activity. There are myriad reasons why an activity may be engaging for one youth but not for another. Thus, these devices and techniques are currently best seen as tools for research, program evaluation, and to inform program, design, but not tools for use in the form of summative assessment of individuals.

To summarize, there are two main takeaways from this chapter. One is that new technologies such as electrodermal activity measurement devices and wearable cameras have the potential to provide some insight into what activities are engaging for youth in makerspaces. Considering that makerspaces are relatively new and highly varied in the experiences that they offer, we stand to gain much if we can specify what maker activities are more or less engaging and for whom they are engaging. The other takeaway relates to differential engagement. If the assumption that moments of psychophysiological arousal are a reasonable proxy for situational engagement, then our tests comparing Dot's and Jane's arousal responses show both similarities and differences in what maker activities were engaging for each youth. Some activities are more engaging for some youth than others. The potential of makerspaces as spaces for interest to develop may lie in their open structure and diversity of experiences they offer. That is, the mix of having the opportunity to work with new tools coupled with opportunities to just watch and listen to others may be part of the promise of making for learning and for developing new interests.

LIMITATIONS

In the previous sections we have argued that understanding engagement in a makerspace in a minute-to-minute scale could be valuable for looking at how engaged learners are with different activities. We proposed using new wearable devices as tools to examine engagement at this time scale. While our method may detect times of heightened engagement, further work must test the validity of electrodermal arousal as a proxy for engagement. This is just an initial analysis of a modest data set in order to develop the method.

We observed unique arousal responses from Jane and Dot using the combination of a psychophysiological measure and first-person point-of-view photos. While we did detect differences between the two, additional work can still be done to examine engagement as a phenomenon of interest. This brief chapter argues that our method is promising, but still a work in progress. We note our tentative assertion of identifying moments of arousal and engagement can still benefit from further justification as this work is more fully developed. Still, we are cautiously optimistic that there may be something to discover through the examination of electrodermal activity data when youth are immersed in learning activities taking place in makerspaces.

ACKNOWLEDGMENT

Portions of this paper were presented previously at the FabLearn 2016 conference. We thank Christopher, the mentors, and the participants for supporting this work. We are grateful to: Joel Drake and Abigail Phillips for their feedback on drafts of this paper, Mac Cannady and Paulette Vincent-Ruz for their help with the engagement survey, and Jeffrey Thayne for assisting with R. This work was supported in part by funding from the National Science Foundation under Grant No. CNS-1623401. The opinions expressed herein are those of the authors and do not necessarily reflect those of the National Science Foundation.

REFERENCES

Andrade, A., & Danish, J. A. (2016). *Using Multimodal Learning Analytics to Model Student Behaviour : A Systematic Analysis of Behavioural Framing*. doi:10.18608/jla.2016.32.14

Arroyo, I., Cooper, D. G., Burleson, W., Woolf, B. P., Muldner, K., & Christopherson, R. (2009). Emotion Sensors Go To School. In V. Dimitrova, R. Mizoguchi, B. Du Boulay, & A. Graesser (Eds.), *Proceedings of the 14th International Conference on Artificial Intelligence in Education* (pp. 17–24). Amsterdam, The Netherlands: IOS Press.

Ben-Eliyahu, A., Moore, D., Bathgate, M., Correnti, R., Randol, S., Dorph, R., & Al, E. (n.d.). Investigating the multidimensionality of engagement: Affective, behavioral, and cognitive engagement in science across multiple days, activities, and contexts. *Contemporary Educational Psychology*. doi:10.1016/j.cedpsych.2018.01.002

Blikstein, P., Gomes, J. S., Akiba, H. T., & Schneider, B. (2016). *The Effect of Highly Scaffolded Versus General Instruction on Students' Exploratory Behavior and Arousal*. Technology, Knowledge and Learning. doi:10.100710758-016-9291-y

Boucsein, W., Fowles, D. C., Grimnes, S., Ben-Shakhar, G., Roth, W. T., Dawson, M. E., & Filion, D. L. (2012). Publication recommendations for electrodermal measurements. *Psychophysiology*, *49*(8), 1017–1034. doi:10.1111/j.1469-8986.2012.01384.x PMID:22680988

Brahms, L., & Crowley, K. (2016). Makeology: Makerspaces as Learning Environments. New York: Routledge.

Calabrese Barton, A., Tan, E., & Greenberg, D. (2016). The makerspace movement: Sites of possibilities for equitable opportunities to engage underrepresented youth in STEM. *Teachers College Record*.

Dailly, S. B., James, M. T., Roy, T., & Darnell, S. S. (2015). EngageMe : Designing a Visualization Tool Utilizing Physiological Feedback to Support Instruction. *Techology, Instruction. Cognition and Learning*, *10*, 107–126.

Fields, D. A., & King, W. L. (2014). So, I think I'm a programmer now": Developing connected learning for adults in a university craft technologies course. In *Learning and becoming in practice: The International Conference of the Learning Sciences* (*Vol. 1*, pp. 927–936). Academic Press.

Hidi, S., & Renninger, K. A. (2006). The Four-Phase Model of Interest Development. *Educational Psychologist*, *41*(2), 111–127. doi:10.120715326985ep4102_4

Kreibig, S. D., Gendolla, G. H. E., & Scherer, K. R. (2012). Goal relevance and goal conduciveness appraisals lead to differential autonomic reactivity in emotional responding to performance feedback. *Biological Psychology*, *91*(3), 365–375. doi:10.1016/j.biopsycho.2012.08.007 PMID:22947258

Lee, V. R., Fischback, L., & Cain, R. (2019). A wearables-based approach to detect and identify momentary engagement in afterschool makerspace programs. *Contemporary Educational Psychology*, *59*, 101789. doi:10.1016/j.cedpsych.2019.101789

Lee, V. R., King, W. L., & Cain, R. (2015). *Grassroots or returning to one's roots? Unpacking the inception of a youth-focused community makerspace*. Paper presented at FabLearn 2015, Stanford, CA.

Martin, L. (2015). The Promise of the Maker Movement for Education. *Journal of Pre-College Engineering Education Research*, *5*(5), 1–30. doi:10.7771/2157-9288.1099

Matsumoto, R. R., Walker, B. B., Walker, J. M., & Hughes, H. C. (1990). Fundamentals of neuroscience. In J. T. Cacioppo & L. G. Tassinary (Eds.), *Principles of pyschophysiology: Physical, social, and inferential elements*. Cambridge University Press.

Pecchinenda, A. (1996). The Affective Significance of Skin Conductance Activity During a Difficult Problem-solving Task. *Cognition and Emotion*, *10*(5), 481–504. doi:10.1080/026999396380123

Poh, M. Z., Swenson, N. C., & Picard, R. W. (2010). A wearable sensor for unobtrusive, long-term assessment of electrodermal activity. *IEEE Transactions on Biomedical Engineering*, *57*(5), 1243–1252. doi:10.1109/TBME.2009.2038487 PMID:20172811

Raskin, D. C. (1973). *Attention and Arousal. Electrodermal Activity in Psychological Research*. Academic Press, Inc. doi:10.1016/B978-0-12-565950-5.50007-7

Renninger, K. A., & Hidi, S. E. (2016). *The Power Of Interest For Motivation And Engagement*. New York: Routledge.

Schneider, B., Krajcik, J., Lavonen, J., Salmela-Aro, K., Broda, M., Spicer, J., ... Viljaranta, J. (2016). Investigating optimal learning moments in U.S. and finnish science classes. *Journal of Research in Science Teaching*, *53*(3), 400–421. doi:10.1002/tea.21306

Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The Challenges of Defining and Measuring Student Engagement in Science. *Educational Psychologist*, *50*(1), 1–13. doi:10.1080/00461520.2014.1002924

Svihla, V. (2015). Making for Engagement, Development and Learning. In Proceedings of FabLearn 2015 (pp. 1–4). Stanford.

Worsley, M., Scherer, S., Morency, L. P., & Blikstein, P. (2015). Exploring Behavior Representation for Learning Analytics. *Proceedings of the 2015 International Conference on Multimodal Interaction (ICMI)*, 251–258.

ADDITIONAL READING

Azevedo, R. (2015). Defining and Measuring Engagement and Learning in Science : Conceptual, Theoretical, Methodological, and Analytical Issues. *Educational Psychologist*, *50*(1), 84–94. doi:10.1080/00461520.2015.1004069

Bevan, B. (2017). The promise and the promises of Making in science education. *Studies in Science Education*, *7267*(April), 1–29. doi:10.1080/03057267.2016.1275380

Braithwaite, J. J., Watson, D. G., Jones, R., & Rowe, M. (2013). A guide for analyzing electrodermal activity (EDA) & skin conductance responses (SCRs) for psychological experiments. *Psychophysiology*, *49*, 1017–1034.

D'Mello, S., Dieterle, E., & Duckworth, A. (2017). Advanced, analytic, automated (AAA) measurement of engagement during learning. *Educational Psychologist*, *52*(2), 104–123. doi:10.1080/00461520.2017.1281747 PMID:29038607

Di Lascio, E., Gashi, S., & Santini, S. (2018). Unobtrusive Assessment of Students' Emotional Engagement During Lectures Using Electrodermal Activity Sensors. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2*(3), 103. 10.1145/3264913

Dougherty, D. (2013). The maker mindset. In M. Honey & D. Kanter (Eds.), *Design, make, play: Growing the next generation ofSTEM innovators* (pp. 7–11). New York, NY: Taylor & Francis.

Ghadiri Khanaposhtani, M., Liu, C. J., Gottesman, B. L., Shepardson, D., & Pijanowski, B. (2018). Evidence that an informal environmental summer camp can contribute to the construction of the conceptual understanding and situational interest of STEM in middle-school youth. *International Journal of Science Education. Part B*, *0*(0), 1–23. doi:10.1080/21548455.2018.1451665

Hsu, Y. C., Baldwin, S., & Ching, Y. H. (2017). Learning through making and maker education. *TechTrends*, *61*(6), 589–594. doi:10.100711528-017-0172-6

Setz, C., Arnrich, B., Schumm, J., La Marca, R., Tröster, G., & Ehlert, U. (2010). Discriminating stress from cognitive load using a wearable eda device. *IEEE Transactions on Information Technology in Biomedicine*, *14*(2), 410–417. doi:10.1109/TITB.2009.2036164 PMID:19906598

Sridhar, P. K., Chan, S. W. T., & Nanayakkara, S. (2018). Going beyond performance scores: Understanding cognitive-affective states in kindergarteners. In *Proceedings of the 17th ACM Conference on Interaction Design and Children* (pp. 253–265). Trondheim, Norway.

KEY TERMS AND DEFINITIONS

EDA: Electrodermal activity (EDA) is a change in a person's skin conductivity due to activation of the sympathetic nervous system. This psychophysiological response of the body causes an increase of sweat production near the hands and feet. EDA can be measured my measuring the resistance to a small current passed through the skin.

Engagement: Engagement with a particular activity or content is a combination of behavioral, cognitive, and affective components. First, behavioral engagement relies on what is observable. Second, cognitive engagement involves actively thinking about the material. Finally, affective engagement includes a positive emotional response to an activity.

High Altitude Balloon: A STEM project where a helium-filled weather balloon transports a payload to the upper reaches of Earth's atmosphere. Payloads typically include cameras and sensors to document the flight. The payload reaches peak altitude when the balloon bursts and falls to the ground beneath a parachute for retrieval.

Makerspace: A makerspace is typically a shared workspace where individuals fabricate objects. Athough these spaces often include digital fabrication tools like 3D printers and laser cutters, hand tools, and repurposed materials are also prominent. Makerspaces promote rapid cycles of design, fabrication, and user testing.

MicroSeimens: The microSeimen (μS) a unit for measuring skin conductivity. It is equal to the inverse of resistance.

RCSC: Relative change in skin conductivity (RCSC) is a method for examining changes in EDA for the purpose of detecting activation of the sympathetic nervous system. Starting with mean levels of skin conductivity 10-second intervals, changes between intervals are calculated to detect periods of rising skin conductivity.

Situational Interest: A situational interest is when environmental factors support a person to be engaged with a particular class of content. This external support results in a person having focused attention on the activity.

ENDNOTES

[1] Youth self-selected into this optional after-school group.

[2] All names have been replaced with pseudonyms to protect participants' identities.

[3] We chose this threshold to build in a margin rather than simply looking at which quantity was greater. We acknowledge that we did not conduct statistical significance testing which could be more robust, but our n within cells was often too small for such testing. Future use of this method would benefit from larger samples and statistical testing of ratios.

[4] As mentioned in the caption of Figure 6, the arousal ratio is represented as a fraction when there are zero images in the aroused or unaroused for a given code. While this arose in cells with low n's, we made the choice to still include them in our analysis. In this case we have a total of 21 arousing images and no unarousing images while the pair were presenting to the group.