Design for Mental Health: How Socio-Technological Processes Mediate Outcome Measures in a Field Study of a Wearable Anxiety App

Alissa N. Antle
School of Interactive Arts and Technology, Simon Fraser University, Surrey, Canada
[aantle@sfu.ca]

Elgin Skye McLaren
School of Interactive Arts and Technology, Simon Fraser University, Surrey, Canada
[elgin-skye_mclaren@sfu.ca]

Holly Fiedler
Education Assistant Department, Langara College, Vancouver, Canada
[hollyfiedler@gmail.com]

Naomi Johnson
Social Work Independent Consultant, Toronto, Canada
[naomi.megan.johnson@gmail.com]

ABSTRACT

Millions of children have challenges with anxiety that negatively impact their development, education and well-being. To address this challenge, we developed version 2.0 of *Mind-Full*, a wearable, mobile neurofeedback system, designed to teach young children to learn to self-regulate anxiety. We present a mixed methods evaluation of a seven week long intervention in schools. We report on a subset of outcome measures related to 10 children's anxiety and stress in the classroom and describe mediating *socio-technological processes* that may have impacted outcomes. Findings showed improvement in children's ability to self-regulate anxiety and reduced cortisol levels for some children. Qualitative findings suggested that children who made multimodal connections during system mediated learning and had teacher support for learning transfer responded well to the intervention. We suggest that framing mental health app design as a distributed, adaptive, socio-technological system enables designers to better meet individual's unique and changing mental health needs.

Author Keywords

Brain-computer interfaces; children; socio-technological studies; emotion-regulation; learning; mental health.

CSS Concepts

• Human-centered computing~Empirical studies in ubiquitous and mobile computing • Social and professional topics~Children

INTRODUCTION

Anxiety disorders are among the most common mental health issues in childhood and the first to emerge – affecting about 18% of children, often lasting into adulthood with devastating social-emotional, health, and economic costs [5,25]. Children who attend schools in lower socio-economic status (SES) areas more frequently have anxiety disorders because they experience ongoing, frequent neurobiological stress responses [24]. These stress responses are correlated with increasing risk of both physical and mental health problems [12]. When these responses occur in periods of rapid brain development, such as when young children first enter schooling, these stress responses can be particularly harmful [12]. Recent therapeutic research suggests that learning to self-regulate anxiety early in life is beneficial and may reduce the negative impact of this mental health challenge on educational, social-emotional, and economic outcomes for children [34]. Self-regulation (SR) training, like mindfulness, and involves learning ways to stay calmly focused and alert, and practicing those techniques until they become automated. Clinical research has shown than bio-feedback and/or neurofeedback may be effective for teaching children how to self-regulate anxiety [13,20,30].

A neurofeedback (NF) system measures, records and provides visual and/or auditory feedback about brain states (e.g. anxiety). Until recently NF systems for self-regulation training were only available to a very limited number of children through clinics and hospitals, and wait times for mental health treatment for children are often 1 year or more [29]. The release of inexpensive consumer grade headsets (e.g. Neurosky, Emotiv and Muse) creates opportunities for NF apps that can be developed for smart phones and tablets, devices children use almost daily in schools. Initial research into school-based NF interventions using consumer grade headsets with children living in poverty [3] and children and youth in Dutch classrooms [32] have shown some positive results and revealed the importance of contextual factors. Our goal for this project was to evaluate a mobile neurofeedback system to help young children in low SES schools learn to self-regulate. To address this goal, we asked: *Can a wearable mobile neurofeedback app for children help them learn to self-regulate anxiety in a SES school setting*? and *What socio-technological processes mediate intervention outcomes?*

In this paper, we present a subset of the results of our mixed methods evaluation (see [4] for more details). Here, we focus on 20 children with high levels of anxiety who used a new

TEI '19, March 17–20, 2019, Tempe, AZ, USA

ACM ISBN 978-1-4503-6196-5/19/03…$15.00
https://doi.org/10.1145/3294109.3295650

version of our *Mind-Full*[1] system. The quantitative results provide evidence that the intervention group improved their ability to self-regulate anxiety at school. We contribute the qualitative findings that identify mediating *socio-technological processes* that impacted outcome measures. We conclude with a recommendation of factors that should be considered in future design and evaluations of socio-emotional and mental health apps for school children.

BACKGROUND

An electroencephalogram (EEG) is a device that uses electrodes placed on the scalp to sense, measure and record the electrical activity of the brain's neurons [18]. As we think, feel, sleep, exercise and learn, electrical activity escapes through our skulls and can be sensed by an EEG device. When this information is displayed on a computer or app screen and the user tries to learn to control their brain states, the system is called a passive EEG-based neurofeedback (NF) system. A single EEG sensor on the left pre-frontal cortex can measure electrical activity, which a computer program can analyze to infer in real time if the EEG wearer has a relaxed or anxious brain state.

Field-based NF Studies of Children and Anxiety

Outside of clinics and research labs, there have been only a few EEG-NF interventions targeting children with anxiety challenges. In a randomized controlled experiment with an active control group with 136 participants (aged 8 to 13), children and youth played one of two platform video games [32]. They played in common classrooms in Dutch schools (7 to19 children per room) for 5 x 1 hour sessions, scheduled twice a week. The intervention group used a NF platform video game called MindLight, developed using principles of cognitive behavioral therapy. The goal was to explicitly expose players to both anxiety producing and relaxing experiences. The control group played Max and the Magic Marker, a puzzle platform video game. Participant and parent ratings of anxiety showed a decrease from pre to post-test, no change at follow-up test points. The methodological design has a large sample size and standardized, reliable and multi-informant behavioral measures. However, the authors suggest that the play context may have impacted results since the children all played in shared rooms (contamination effect). The authors also suggest that positive results might be due to non-specific factors such as being primed to expect an improvement in anxiety. Lastly, they acknowledge that without a non-intervention group it was difficult to determine if improvements were greater than without any intervention.

In a randomized controlled experiment with a waitlist control group with 21 girls (aged 5 to 11), children completed 24 NF sessions over 6 weeks with their counsellor at a school for girls living in poverty in Nepal [3]. The Mind-Full NF system had three simple games, designed based on principles described in [2] including the use of embodied metaphor based interaction that linked body enactments with changes in brain states represented in ways children could understand from everyday activities. The goal was to teach the girls to learn to use breathing and body relaxation techniques to SR anxiety and transfer those skills to the classroom and playground of the school. Counsellor and teacher ratings of anxiety showed an increase in ability to SR anxiety in a variety of contexts for the intervention from pre to post-test, no change at follow-up test points. The waitlist group later completed the intervention and pre-post ratings showed a similar pattern. Interviews with counsellors and teachers highlighted system-based factors including the benefits of repetitive practice, the laddering structure of the games and the ability to adjust the game difficulty in real time.

Interviews with school staff combined with researcher observations revealed several contextual factors that may have been related to beneficial outcomes [3]. First, counsellors reported that seeing the children's brain states in real time helped them better understand and treat the challenges faced by each child, leading to changes in their counselling approaches. Second, meetings between researchers and school staff helped create a common language for talking about the challenges faced by the children. In combination with the access to brain data over the internet, this may have enabled remote western therapists to better coach the Nepalese counsellors. Third, the children's attendance rates improved over the study period, in part because parents may have been motivated to ensure their children were exposed to western technologies. Lastly, counsellors reported that after each session the girls were often more relaxed and willing to talk about their anxiety, which likely improved the rapport between them and/or the quality of the one-on-one counselling sessions.

To the best of our knowledge, there are no other field studies of NF-based interventions for anxiety and children.

Field-Based and Contextual Methodological Challenges

For further methodological guidance we can look to other field-based NF studies targeting children and other brain processes (e.g. attention) [8,9,17,22,37]. Claims from these studies are often limited because study designs have involved small sample sizes, non-randomized designs, and inadequate control conditions [8]. However, we note that when working in lower SES contexts, an active (placebo) control group design may be ethically unacceptable; due to committing limited resources to a placebo-based intervention [1]. Another issue noted is that some studies do not assess a child's ability to transfer SR skills into their everyday lives or assess maintenance over time [10,36]. Lastly, a common discussion point is that it remains unknown if NF interventions that show behavioral improvements also help children improve regulation of neurological or physiological processes (e.g. stress response) [38]. We suggest that what is needed is a field study design in line with the practices noted above but with a waitlist control group and measures of

[1] http://www.mindfullapp.ca/

transfer and maintenance of SR of anxiety, behaviors related to anxiety and physiological stress responses.

MIND-FULL: MOBILE EEG NEUROFEEDBACK APP

Our new *Mind-Full* system is comprised of six relaxation games that are played with an EEG headset and a tablet, and a second connected tablet that runs a calibration module. Each game is based on familiar, everyday activities and actions which, when learned, can elicit behaviors which in turn result in desired brain states related to relaxation. The games are designed to help children learn and practice SR of anxiety. The calibration module enables a facilitator to adjust brain state targets and game difficulty during real-time gameplay.

The system uses a BrainLink Pro[2] EEG headset to wirelessly transmit brainwave data into a Unity 3D game engine running on an Android interactive tablet. Brainlink is a commercially available EEG headset that uses the Neurosky chip and only one dry electrode to record electrical activity in the left pre-frontal cortex. Compared to most EEG units, it is robust, non-invasive, adjustable, and easy to wear due to the single dry electrode. The headset sends data via Bluetooth to apps in Unity 3D where it is analyzed and used to control game mechanics, and interface and output elements.

Each game was designed to help children learn how to SR anxiety by enacting actions that calm their body, which in turn calms their brain. The games were all designed using embodied metaphor theory, as outlined in [2], to help young children (aged 5 to 8) understand the relationship between their body and brain as they learn to SR anxiety. We created six games to address concerns of reduced engagement and enjoyment over longer interventions cited in [19].

Relax Games: In three similar games children learn to relax by controlling their breathing. For example, in the indigenous themed Salmon game we represent relaxation using a metaphor – exhaling softly is calming. The game begins with an animation in which a child blows softly on an ember from a lightning strike to ignite a flame. The flame then grows and cooks (smokes) a salmon, which is an indigenous tradition familiar to children in our geographical area. (Figure 1). After the animation plays, it is the user's turn. Like the animated girl, the user softly blows on the flame image, which relaxes her body, thus shifting her brain state toward a relaxed state (lower alpha brainwave frequency). The EEG sensor uses a proprietary algorithm to determine a "relax" index (0-100) that is sent through Bluetooth to the Unity 3D game. When the index passes a threshold for a set amount of time, the game program triggers an animation of the flame growing larger and the salmon blackening. To add a playful twist, the cooked salmon is then stolen by a pesky raven. This sequence earns the user a token stored in a container (Figure 1 left side). If the user is not able to relax the flame goes out and they must start over. Each game is complete when five tokens are earned. The other two relax games are similar involving controlled and gentle breathing.

Sustained Relaxation Games: In three other games children learn how to sustain a relaxed brain state by relaxing and then remaining in a relaxed body state. For example, in the Monster game the user must relax their body as they stay quiet and still in the forest to catch the monster who is chasing a squirrel over a leaf trap (Figure 2). If the user does not remain relaxed the monster runs away into the forest and they must start over. All three games involve sustained relaxation based on cues to enact a relaxed, still body state.

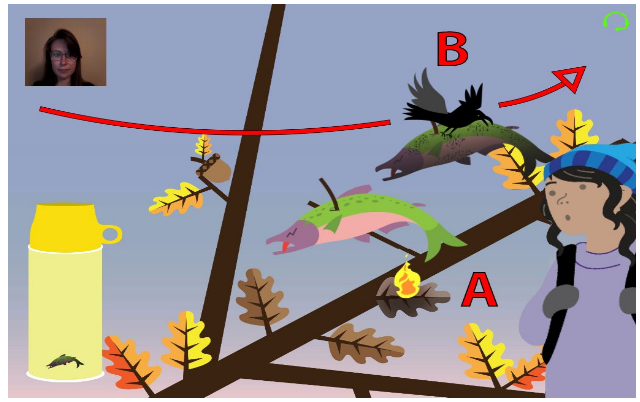

Figure 1. Salmon game. A: Blow softly to grow flame. B: Smoked salmon is then stolen by raven.

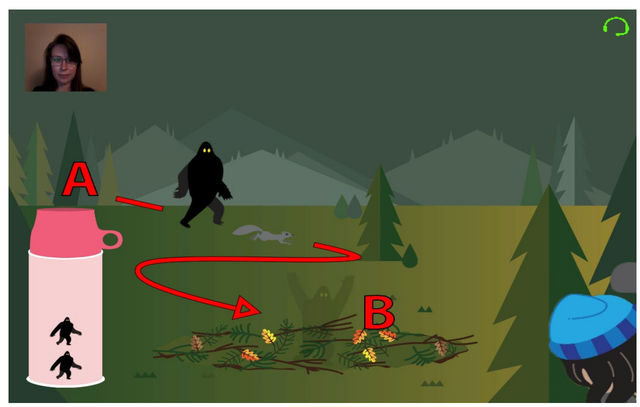

Figure 2. Monster game. A: Relax body to hide so monster chases squirrel. B: Sustain relaxation so monster falls into trap.

The calibration module enables the facilitator to view and adjust brain state targets in real time to accommodate for individual differences (Figure 3). Most NF systems require calibration before use through training. However, children in our population cannot reliably create a relaxed brain state for long enough to train the system. To account for this challenge we developed a real time calibration module. The system is preset with a typical "relax" index value of 40 on a 0-100 scale. However, this can be adjusted in real time. For example, if the facilator sees that a child appears relaxed and

[2] http://store.neurosky.com/collections/brainlink-pro

is breathing softly and the module shows a state just less than 40, they can adjust the target to 35 to accommodate this individual variation from the typical target. As the child learns to relax they can also make this target higher/harder (e.g. 60/100). Lastly, the facilitator can also adjust the amount of time that a brain state must be held to earn a token, again to make each game easier or harder - to accommodate both improvement over time and also variations that children have day to day in their ability to relax depending on events in their lives (e.g. a stressful day at school).

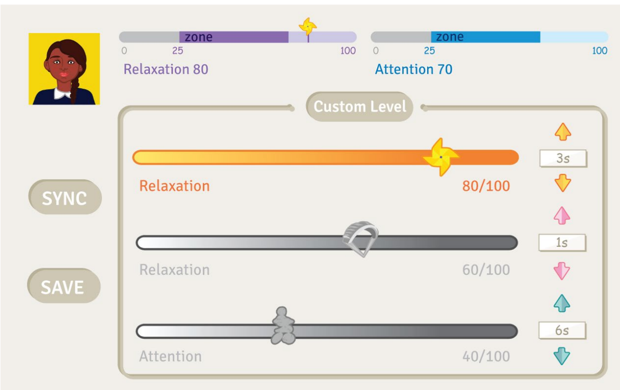

Figure 3. Calibration module enables real time view and change of user's brain state and hold time parameters for games.

METHODOLOGY

We conducted a 2x2 mixed methods evaluation with 20 children (aged 5 to 8 years) who had high levels of anxiety. Children were split into two equivalent groups (intervention, control). All children were assessed at pre and post-test points, and the intervention group was also assessed 2 months later at a follow-up point. The independent variables are group and test point. The control group acts as a comparator, accounting for children's social-emotional learning at school or home as well as any other maturation or developmental effects. At a later date, for ethical reasons, the control group also did the intervention (data not reported here). Our research questions and outcome measures (dependent variables) are described below. We examined both behavioral and physiological indicators of anxiety. We also collected data to understand the impact of socio-technological processes on outcomes that may occur in a complex field environment.

Participants

Thirty-three children were identified by school principals, and/or head teacher and classroom teachers at two elementary schools. The primary criteria for selection was reported severe level of anxiety. The schools served diverse, multi-cultural populations, including children from low SES households, recent immigrants and refugees. Most of the children had suffered from trauma due to family dysfunction, refugee status, marital break-ups and/or other events associated with low SES.

Of the 33 children identified, 21 were identified by teachers as having severe anxiety and age appropriate cognitive functioning. Twenty children and their parents provided assent/consent to participate. We formed 10 matched pairs of children based on grade, gender, anxiety challenges, and personal characteristics. One child from each pair was randomly assigned to the intervention group and one to the waitlist control group. The end result was 14 males, 6 females split evenly across groups and both schools. Children's teachers also participated in the study (n=7), all female and each with five+ years of teaching experience. The children's ages ranged from 5 to 8 years old at the beginning of the study. All children had good mastery of the English language. We verified group equivalence in terms of ability to SR anxiety using the Calm teacher survey (see below). Ethics was approved by Simon Fraser University and SD41.

Procedure

The intervention group participated in 18 facilitated sessions with our system over seven weeks. The control group did not use the system over this period. The intervention was delivered in 15-20 minute practice sessions facilitated by a recently graduated social worker. An educational assistant (EA) trainee who was part of the research team was also on site to take notes, manage the schedule and data collection. Sessions were delivered over six weeks with a frequency of three times per week, and the seventh week was used for make-up sessions. The length, frequency and number of sessions was based on other successful NF studies with children (e.g. [7,9,19,32]) and adapted to class schedules.

Each child was introduced to our system using a scripted protocol. Teachers were also introduced to the intervention and trained to use simple posters and phrases taken from the intervention to support children to transfer SR skills from the games into the classroom. We did teacher training in 10-15 minute sessions with teachers in small groups, or individually as scheduling allowed.

Figure 4. Child using system in study setting.

Each child's sessions were facilitated with a scripted protocol designed to identify daily stresses, encourage the child to focus on practicing rather than succeeding at playing the games, and to use their SR skills at school. Facilitation also involved helping children understand the relationship between their body and brain (e.g. breathing calms the brain, relaxing the body from feet all the way up to head helps body stay relaxed). Each session took place in a small quiet room in each school with chairs, table and a couch (Figure 4). Each session was structured beginning with a discussion about how the child was doing on that day and if/how they had used their SR skills since the last session. Then the facilitator explained/reminded the child how to play the relax game using deep, slow breathing and the child played that game until they received five tokens. The facilitator next explained/reminded the child how to play the sustained relaxation game by relaxing their whole body, and the child played that game until they received five tokens. During gameplay, if needed, the facilitator used the calibration module to adjust brain state targets or game difficulty. Children played each of the three relax and three sustained relaxation games for two of the six weeks, in the same order. Sessions ended with a discussion of how the child could use SR skills in situations at school and home.

Research Questions, Data Collection and Analysis

We collected session data with structured facilitator observational notes to record the number and duration of sessions, and number of tokens for each game to determine if children were able to complete the intervention. We assessed all the children (intervention and control) at pre and post-test points using numeric surveys (teacher), interviews (facilitator, teacher), and physiological cortisol tests (child) (all described below). The intervention group was again assessed (surveys, interviews) at a 2 month follow-up.

RQ1 (SR Anxiety): *Do children in the intervention group more effectively self-regulate anxiety than those in the control group in a school environment?*
To address this question we collected data for the *SR Anxiety* construct, which is children's ability to SR anxiety, in both groups at pre and post-test points, and for the intervention group again at two month follow-up. Our measurement instrument was a variant of the Calm survey reported in [4]. Using the Calm survey, *SR Anxiety* is operationalized by a teacher rating a child's ability to SR anxiety with support in different school situations using a five point rating scale, ranging from: Cannot do this at all even with support (1) to Can do this mostly by themselves (5). In previous work the Calm survey was shown to be reliable (Cronbach's $\alpha > 0.70$) and sensitive to change over time [4]. We adapted the Calm survey to reflect typical contexts encountered in the school by the children in our study. Sample questions included: *Child can calm down when they are having difficulty with some schoolwork* and *Child can calm down after a stressful event in the playground.* Since we adapted the survey we assessed its reliability using Cronbach's alpha in SPSS. For Cronbach's alpha values of 0.8 and above, we included all questions. When values were less than 0.8, if deleting any items could increase the overall reliability above 0.8 we removed them to improve the reliability of the construct measure as is standard practice in survey development [6].

RQ2 (Anxious Behaviors): *Do children in the intervention group have a lower frequency of anxiety related behaviors than those in the control group in a school environment?*
To address this question we collected data using the *Anxious Behaviors* construct, which is the frequency of a child's anxious behaviors at school in both groups at pre and post-test points, and for the intervention group at two month follow-up. Our measurement instrument was the Anxiety subscale of the Teacher Behavior Assessment System (BASC-3) survey for children [31]. The BASC-3 has excellent test score reliability (mid .80s or higher) [31]. The teacher survey contains statements with descriptors of 16 behaviors related to anxiety at school and a four point frequency rating scale, ranging from never (1) to almost always (4). Sample statements include: Has panic attacks; Appears tense; and Overreacts to stressful situations.

Analysis of all survey data was done using SPSS to check for assumptions of normality (Shapiro-Wilks' test) and equal variance (Levene's test). We generated descriptive statistics (mean, standard deviation). We ran mixed 2x2 (repeated measures) ANOVAs with independent variables group and test point and Calm and BASC ratings as separate dependent variables. We would expect no between group differences in Calm scores at pretest since the groups were chosen to be equivalent. We expect high variability with a population sample that has severe mental health challenges so we examined both between and within group differences from pre to post-test. We used a repeated measures t-test on the intervention group data to compare the survey ratings from post to follow-up test points. We used an alpha level of .05 for all statistical tests.

RQ3 (Stress Level): *Do children in the intervention group have a lower stress level than those in the control group in a school environment?* To address this question we measured the *Stress* construct by testing salivary cortisol levels at pre and post-test points. A salivary cortisol test provides an objective, biological indicator of stress and has been validated with young children (e.g. [23,26]). We took salivary cortisol swabs following standard procedures for testing with children (e.g. [14,15]). For example, we collected samples at set times of day (9:15 am, 11:30 am, 2:30 pm) and had children avoid vigorous exercise or food intake prior to sampling. Samples were couriered to the Clemens Kirschbaum laboratory at the Dresden University of Technology in Germany for analysis. The lab uses a standard procedure as reported in [28].

Analysis of processed cortisol test data indicated that there were missing values (either no test or empty vial), significant events that impacted results (e.g. crying, sick), and cases where the test protocol had not been followed (e.g. child had just exercised, taken asthma or cold medication). These values were excluded. Exploration of remaining values showed high individual variability, which has been found to be the case for children in other studies [35]. As a result, we conducted a frequency analysis of individual cases. Previous research has established norms for cortisol levels for 5 to 10 year olds throughout the day in the 5^{th}, 50^{th} and 95^{th} percentiles [27]. Interpolation between the 50^{th} and 95^{th}

percentiles is nonlinear and would require complex modelling, so we chose the 50th percentile as a norm. Using this norm data we calculated thresholds for males and females for these cortisol-related dependent variables: 1) morning, 2) average daily and 3) daily decline based on our sample times. For each child and test point we determined if their cortisol variables were above the 50th percentile.

RQ4: *What socio-technological processes may have influenced the intervention outcomes?* In order to address this question researchers, the EA and the facilitator took observational notes for sessions. We conducted three open-ended interviews with the facilitator and EA during the intervention, recorded on audio and with notes. We asked them to describe what they thought was working and not working in the intervention. We asked them about how each child was progressing and what was helping or not helping each child. We also interviewed the teacher of each child at pre, post and follow-up test points and asked questions about each child's ability to SR anxiety (e.g. *Does this child have challenges with anxiety or stress? Describe last month*) and about if they were using cues from the intervention to help the child, and if so, how that was working or not working (e.g. *If you referred to the posters or our system in any way during class time to remind this child to calm down, can you describe what happened?).*

Analysis of interview and observational notes included two researchers separately looking for common or interesting themes that highlighted how socio-technical factors may have impacted or mediated intervention outcomes at the group and individual level. We did member checking with resulting themes to improve validity.

RESULTS

Analysis of detailed session notes taken by the facilitator revealed that all 10 children in the intervention group were able to complete 18 sessions. Each session included about 10 minutes of gameplay in the relax and sustained relaxation games. Children were able to achieve five tokens in each game in each session as well as talk about their anxiety and SR skills to some extent, varying by child.

RQ1: *SR Anxiety* Descriptive statistics comparing teachers' ratings on the Calm survey for each group at each test point are shown in Table 1. We ran a two-way mixed ANOVA on a sample of 19 participants (one child moved before post-test survey) to examine the effect of group and test points on average Calm score. Results indicated a significant interaction between the effects of group and test points on Calm score, $F(2,18) = 4.868$, $p = .041$ with a small effect size ($\eta^2_p = .222$). Simple main effects analysis showed no significant difference at pre or post-test between groups. Within group analysis showed that the intervention group improved significantly ($p = .010$), from pre to post-test. A t-test showed that this improvement was maintained ($p = 1.00$) at the follow-up test. In summary, the intervention group significantly improved in their ability to SR anxiety at school and gains were maintained over time.

Measure	Group	Mean	Std. Dev.	N
Teacher *Calm @* Pretest	Interv'n	2.88	.945	10
	Control	3.11	.574	9
Teacher *Calm @* Post-test	Interv'n	3.43	.834	10
	Control	3.06	.990	9
Teacher *Calm @* Follow-up	Interv'n	3.43	.858	10
	Control	N/A	N/A	0

Table 1. Teacher rating (1-5) ability to self-regulate anxiety.

RQ2: Anxious Behaviors Descriptive statistics comparing teacher ratings of the BASC Anxiety subscale for each group at each test point are shown Table 2. Two-way mixed ANOVA results for average Anxiety score showed no significant effect of group on Anxiety $F(1,18) = .002$, $p = .969$) at either pre or post-test. In summary, the groups were equivalent at pretest as expected, but teachers did not see a change in the numbers or kinds of anxious behaviors over the course of the study between or within groups.

Measure	Group	Mean	Std. Dev.	N
Teacher *Anxiety @* Pretest	Interv'n	2.05	.272	10
	Control	2.00	.601	9
Teacher *Anxiety @* Post-test	Interv'n	1.99	.301	10
	Control	2.06	.571	9
Teacher *Anxiety @* Follow-up	Interv'n	2.05	.319	10
	Control	N/A	N/A	0

Table 2. Teacher BASC-3 rating (1-4) Anxiety.

RQ3. Stress Level Case by case analysis of viable salivary cortisol samples indicated that at pretest **morning cortisol levels** were above the 50th percentile for over half the children in both groups (Table 3). At post-test this count dropped for the intervention group (4/10) and increased for the control group (8/10). At pre and post-test points, the **average daily cortisol** was above 50th percentile for both groups. At pretest, the expected **decline** in cortisol level over the day was only seen for one child. At post-test just over half (4/7) of the intervention group had the expected decline and only one in the control group. In summary, cortisol tests showed high variability, which impacted the average measure, and indicated reduced physiological stress in the morning and daily decline for the intervention group.

Group	Morning		Daily Ave		Decline	
	pre	post	pre	post	pre	post
Interv'n	6/9	4/10	7/8	4/4	1/8	4/7
Control	5/7	8/10	8/8	8/8	0/8	1/9

Table 3. Count of cortisol variables above 50th percentile.

RQ4: Socio-Technological Processes Three common themes that showed how socio-technological processes may have impacted the outcomes reported above emerged from thematic analysis of observations and interviews.

System Mediated Multi-Modal Connections. We noticed that the system was used as a reference for different modalities of communication to help children learn and apply the mind-body relationship in SR. For example, the facilitator noted that she often spent time before and during each game verbally describing and physically demonstrating how a child could use their body to SR while they played the games. We observed that the game animations of body-based enactment linked to brain state feedback helped children translate from words and physical descriptions to physical enactment of actions that helped them feel calmer. The facilitator also noted that some children reported that they had used SR strategies that they had discussed and physically practiced with the system later during an upsetting event (e.g. doing deep breathing and relaxing fists before getting a shot at the doctor's office, or sitting down and breathing slowly after tripping down the stairs). Conversely the facilitator noted that three (of 10) children, who struggled with the games or did not improve over time had trouble conceptually understanding and talking about the mind-body link. We suggest that the ways that the system mediated verbal and physical communication to make it concrete and experiential for children is an important element of the intervention.

Teachers also reported evidence that the system helped some but not all children learn both words and actions related to SR. For example, at post-test, a teacher reported that in a stressful situation with peers a child was more able than she had previously been to recognize and articulate how her body was feeling, and that she could transition and move through activities more smoothly when reminded to do breathing as she had in the sessions. At post-test teachers reported other effects related to children beginning to understand their mind-body relations and be able to talk about their emotions. One teacher said, "It's neat to see them understand how to regulate themselves." One teacher reported that a parent had noticed changes, "His mother - she said at home she was noticing that he can calm his emotions better, describe his emotions more and come up with strategies to calm himself." The teacher reported no change for other children, several of whom were not able to talk about what they felt nor enact SR strategies in the classroom even with support. It seems that the system provided a focus for only some children to learn to talk about body and mind connections, practice SR in ways that they could concretely experience these relations and subsequently verbally articulate how they were feeling.

Customization, Adaptation and Rapport: The facilitator reported that her ability to customize the brain state targets and difficulty of games in real time, without the child knowing helped her help the children complete the games. Sometimes this meant making a game slightly easier as time was running out. Other times it meant making games harder to keep the challenge level high. The facilitator also noted that for some children, who struggled, on some days she had to make the targets very low or the child would give up. She suggested a lack of attention and/or motivation rather than ability might be underlying some of these situations. It was also noted that sometimes this occurred when a child was having a very difficult day at school. Overall, the facilitator used this feature more extensively than we had envisioned. She also reported creating a light-hearted narrative around one of the games involving the squirrel and the monster because some of the children found the scenario frightening. Other children, mainly older boys, liked the monster character. Teachers also mentioned the need for customized and adapted approaches to learning SR. One teacher said, "It's good for them to learn different strategies because sometimes one strategy doesn't work."

An interesting related finding was that the facilitator reported that some of the children were not improving in gameplay but that their behavior in the sessions was improving. We suggest that this may be, in part, a result of the strong rapport established between the social worker and the children, which was commented on by many school staff, the children themselves and observed by the researchers. We also observed that strong rapport helped the facilitator learn worked for each child, trying different variations to respond to the child's needs, which often varied from day to day. The teachers also consistently said that the children enjoyed coming to the sessions and having one-on-one time with the facilitator. Based on these results we suggest that the ability to adapt to individual needs requires that a person who has strong rapport with the child rather than an algorithm drives the customization. In this light, we view our system, including child, facilitator, teacher and technology as a distributed learning system that adapts in real time to meet the needs of each child at each moment. How well the system components function together may dictate how effectively the child can learn to SR.

Transfer: Looking Beyond the Technology. Results from facilitator and teacher interviews provide insight into the ways that cueing students about the intervention helped them transfer their newly learned ability to SR into the classroom. Several teachers reported that reminding students to do breathing like in the relax games helped them remember and try to calm down. For example, one teacher said "If I can get to him before – before it escalates – to do his breathing it works. One time he was just about in tears and I said try your breathing that you do with the iPad and he did a deep breath and was able to do it, it was cool, he de-escalated." Similarly, another teacher reported, "It's getting better. His episodes are shorter and less frequent. Usually his head is against the wall and there's tears and I remind him to blow … he seems to get control faster." These findings highlight the importance of the teacher's attentiveness to the emotional state of the child, and that the timing of interaction with the child is critical to the child successfully applying skills learned through technology in the social context of a classroom. Conversely, teachers noted that if they didn't notice or intervened too late, it was likely that the child would not be able to SR on their own. When we looked at individual pre-post survey scores, we noted that two (of 10) children who did not improve on *SR Anxiety* were in a classroom where the teacher did not report using cues to help the children practice SR.

DISCUSSION

This is the first study to show that it is viable for a non-clinically trained facilitator to use an EEG neurofeedback brain computer system in a school setting with young children. All 10 of the children in the intervention group were able to complete the 18 sessions. There was a significant interaction effect between time and group. Within group simple effects showed significant gains for *SR Anxiety*, which were maintained, as was seen in [3]. However, between group effects at post-test, were non-significant. An examination of the means showed little change in the control group means, a gain for the intervention group and high variability. The small sample size and non-blind raters may have impacted these results. In addition, other NF studies have identified responders and non-responders [10], and our analysis of facilitator notes and interview suggest that our *SR Anxiety* results were impacted by three non-responders. Teacher ratings of anxious behaviors (BASC) showed no change, which is in contrast to [32] in which both groups of older children were rated as improving. Stress as measured by cortisol tests showed gains for the intervention group for morning and daily decline after the intervention. We suggest that this may be in part from learning to SR, but is also likely in part, due to the impact of the strong rapport the children had with the facilitator and the break from classroom routines enabled by their 18 1:1 sessions. Follow-up analysis of stress would help disambiguate these factors.

We also found that socio-technological process factors that impacted both positive and negative outcomes included the ability of the children to learn to connect words, feelings, body sensations with self-regulation skills through NF and transfer of that learning to new situations with explicit and well timed support from teachers. Children that struggled conceptually or were less motivated or attentive in sessions often were those who had difficulty self-regulating in class. The rapport with the facilitator and her ability to adapt the system and protocol to meet individual needs was also likely critical to success. These findings are in line with other learning sciences research. The theory of mutual adaptation that posits interdependence in early learning can support children to be less dependent on their immediate environment (in our case the session) and more adaptive when they confront new environments (SR in the classroom) [33]. Similarly, much of the literature in experiential learning highlights the inter-relations of cognitive, body-based and affective modalities in learning and the need for concrete opportunities for transfer (e.g. [21]). Research on socio-emotional learning through games stress the importance of motivation and facilitator rapport [16]. Other research suggests that it is critical to train teachers to conduct transfer activities [36], which should ideally leverage existing classroom practices related to social-emotional learning adapted to experiences and language taken from the intervention. Children need considerable support to transfer what they learn in sessions and this component should be enhanced in future study designs.

Based on these preliminary findings we suggest a framework for mental health app development through the lens of distributed and adaptive systems where consideration is given to the inter-relations between people (child, teacher, facilitator), the activities (sessions, classroom) and technology (system, system as reference) all situated in a specific environment (school). Consideration should be given to the inter-relations between distributed components to ensure connection can be made through multimodal forms of communication and learning with resources dedicated to transfer activities. The system must be adaptable over time to respond to individual and changing needs of children with mental health challenges.

In contrast to the tradition of using clinical trials or controlled experiments to assess digital health interventions, our field-based results highlighted the importance of considering mediating socio-technological factors. Given the individual variability in our population sample, we suggest that in future evaluations using within group designs such as N-of-1 studies in which children are their own comparators, would enable researchers to look at individual change. In our study, the way our facilitator customized the intervention to meet the needs of individual children is a confounding factor that limits claims but suggests to us a more ethically sound approach to working with such vulnerable populations. We suggest purposely creating technologies and interventions that can be adapted throughout the study to mitigate risk and improve outcomes for high risk individuals.

CONCLUSION

Despite the challenges of working with this population, we targeted them because we know that unmitigated early stress experiences can permanently impact the neurobiological systems involved in anxiety and stress regulation [12]. Using NF with commercially available EEG headsets on mobile devices children have, are familiar with, and have access to may provide a way to provide mental health support to many more children than can be supported in clinics or hospitals. Our research suggests that a mobile wearable NF app may be used in schools to help teach SR. Based on our work and that of others (e.g. [3,11,32]) we suggest taking a socio-technological approach to design and evaluation so that we move beyond only looking at outcomes measures to ensure we can effectively address the growing gap between mental health challenges and ehealth support services. We see this study as one small step in this emerging design space.

ACKNOWLEDGEMENTS

Thanks to NSERC, Leslie Chesick, the Burnaby School District, Mia Cole for wrangling cortisol samples, and as always, special thanks to all children and their teachers.

REFERENCES

[1] Alissa N. Antle. 2017. The ethics of doing research with vulnerable populations. *ACM Interactions* 24, 6:74-77.

[2] Alissa N. Antle, Leslie Chesick, and Elgin Sky McLaren. 2018. Opening up the design space of neurofeedback brain-computer interfaces for children. *ACM Transactions on Computer-Human Interaction* 24, 6: 1–38.

[3] Alissa N. Antle, Leslie Chesick, Srilekha Kirshnamachari Sridharan, and Emily Cramer. 2018. East meets west: A mobile brain-computer system that helps children living in poverty learn to self-regulate. *Personal and Ubiquitous Computing*: 1–28. https://doi.org/10.1007/s00779-018-1166-x

[4] Alissa N. Antle, Elgin Skye McLaren, Holly Fiedler, and Naomi Johnston. 2019. Evaluating the impact of a mobile neurofeedback app for young children at school and home. In *Proceedings of the Conference on Human Factors in Computing Systems* (CHI '19), in press.

[5] Katja Beesdo, Susanne Knappe, and Daniel S. Pine. 2009. Anxiety and anxiety disorders in children and adolescents: Developmental issues and implications for DSM-V. *Psychiatric Clinics of North America* 32, 3: 483-524.

[6] Ellen Drost. 2011. Validity and reliability in social science research. *Education Research & Perspectives* 38, 1: 105–123.

[7] Holger Gevensleben, Birgit Holl, Björn Albrecht, Claudia Vogel, Dieter Schlamp, Oliver Kratz, Petra Studer, Aribert Rothenberger, Gunther H. Moll, and Hartmut Heinrich. 2009. Is neurofeedback an efficacious treatment for ADHD? A randomised controlled clinical trial. *Journal of Child Psychology and Psychiatry* 50, 7: 780-789. https://doi.org/10.1111/j.1469-7610.2008.02033.x

[8] Holger Gevensleben, Maike Kleemeyer, Lillian Geza Rothenberger, Petra Studer, Andrea Flaig-Röhr, Gunther H. Moll, Aribert Rothenberger, and Hartmut Heinrich. 2014. Neurofeedback in ADHD: Further pieces of the puzzle. *Brain Topography* 27, 1: 20-32. https://doi.org/10.1007/s10548-013-0285-y

[9] John. H. Gruzelier, M. Foks, T. Steffert, M. J. -L. Chen, and T. Ros. 2014. Beneficial outcome from EEG-neurofeedback on creative music performance, attention and well-being in school children. *Biological Psychology* 95: 86-95. https://doi.org/10.1016/j.biopsycho.2013.04.005

[10] John H. Gruzelier. 2014. EEG-neurofeedback for optimising performance. III: A review of methodological and theoretical considerations. *Neuroscience & Biobehavioral Reviews 44*: 159-182. https://doi.org/10.1016/j.neubiorev.2014.03.015

[11] John H. Gruzelier. 2014. EEG-neurofeedback for optimising performance. I: A review of cognitive and affective outcome in healthy participants. *Neuroscience & Biobehavioral Reviews* 44: 124-141. https://doi.org/10.1016/j.neubiorev.2013.09.015

[12] Megan Gunnar and Karina Quevedo. 2007. The neurobiology of stress and development. *Annual Review of Psychology* 58, 1: 145-173. https://doi.org/10.1146/annurev.psych.58.110405.085605

[13] D. Corydon Hammond. 2005. Neurofeedback with anxiety and affective disorders. *Child and Adolescent Psychiatric Clinics of North America* 14, 1: 105-123.

[14] Kirsten Hanrahan, Ann Marie McCarthy, Charmaine Kleiber, Susan Lutgendorf, and Eva Tsalikian. 2006. Strategies for salivary cortisol collection and analysis in research with children. *Applied Nursing Research* 19, 2: 95-101.

[15] Amanda G. Harmon, Leah C. Hibel, Olga Rumyantseva, and Douglas A. Granger. 2007. Measuring salivary cortisol in studies of child development: Watch out—what goes in may not come out of saliva collection devices. *Developmental Psychobiology* 49, 5: 495-500.

[16] Robyn Hromek and Sue Roffey. 2009. Promoting social and emotional learning with games: "It's fun and we learn things." *Simulation & Gaming* 40, 5: 626-644. https://doi.org/10.1177/1046878109333793

[17] Jin Huang, Chun Yu, Yuntao Wang, Yuhang Zhao, Siqi Liu, Chou Mo, Jie Liu, Lie Zhang, and Yuanchun Shi. 2014. FOCUS: Enhancing children's engagement in reading by using contextual BCI training sessions. In *Proceedings of the Conference on Human Factors in Computing Systems* (CHI '14), 1905-1908.

[18] Alice F. Jackson and Donald J. Bolger. 2014. The neurophysiological bases of EEG and EEG measurement: A review for the rest of us. *Psychophysiology* 51, 11: 1061-1071. https://doi.org/10.1111/psyp.12283

[19] Stuart J. Johnstone, Steven J. Roodenrys, Kirsten Johnson, Rebecca Bonfield, and Susan J. Bennett. 2017. Game-based combined cognitive and neurofeedback training using Focus Pocus reduces symptom severity in children with diagnosed AD/HD and subclinical AD/HD. *International Journal of Psychophysiology* 116: 32-44.

[20] Michele Knox, J. Lentini, T.S. Cummings, A. McGrady, K. Whearty, and L. Sancrant. 2011. Game-based biofeedback for paediatric anxiety and depression. *Mental Health in Family Medicine* 8, 3: 195.

[21] David. Kolb. 1984. *Experiential Learning: Experience as the Source of Learning and Development*. Prentice-Hall, Englewood Cliffs, NJ, USA.

[22] Choon Guan Lim, Tih-Shih Lee, Cuntai Guan, Sheng Fung DS, Yin Bun Cheung, S. S. Teng, Haihong Zhang, and K. Ranga Krishnan. 2017. Effectiveness of a brain-computer interface based programme for the treatment of ADHD: a pilot study. *Psychopharmacol Bull*. 43, 1: 73-82.

[23] Jared A. Lisonbee, Jacquelyn Mize, Amie Lapp Payne, and Douglas A. Granger. 2008. Children's cortisol and the quality of teacher–child relationships in child care. *Child Development* 79, 6: 1818-1832.

[24] Sonia. J. Lupien, Suzanne King, Michael J. Meaney, and Bruce S. Mcewen. 2001. Can poverty get under your skin? Basal cortisol levels and cognitive function in children from low and high socioeconomic status. *Development and Psychopathology* 13, 3: 653-676.

[25] John S. March. 2011. Looking to the future of research in pediatric anxiety disorders. *Depression and Anxiety* 28, 1: 88-98. https://doi.org/10.1002/da.20754

[26] Ann Marie McCarthy, Kirsten Hanrahan, Charmaine Kleiber, M. Bridget Zimmerman, Susan Lutgendorf, and Eva Tsalikian. 2009. Normative salivary cortisol values and responsivity in children. *Applied Nursing Research : ANR* 22, 1: 54. https://doi.org/10.1016/j.apnr.2007.04.009

[27] Robert Miller, Tobias Stalder, Marc Jarczok, David M. Almeida, Ellena Badrick, Meike Bartels, Dorret I. Boomsma, Christopher L. Coe, Marieke C. J. Dekker, Bonny Donzella, Joachim E. Fischer, Megan R. Gunnar, Meena Kumari, Florian Lederbogen, Christine Power, Carol D. Ryff, S. V. Subramanian, Henning Tiemeier, Sarah E. Watamura, and Clemens Kirschbaum. 2016. The CIRCORT database: Reference ranges and seasonal changes in diurnal salivary cortisol derived from a meta-dataset comprised of 15 field studies. *Psychoneuroendocrinology* 73: 16-23. https://doi.org/10.1016/j.psyneuen.2016.07.201

[28] Eva Oberle and Kimberly A. Schonert-Reichl. 2016. Stress contagion in the classroom? The link between classroom teacher burnout and morning cortisol in elementary school students. *Social Science & Medicine* 159: 30-37.

[29] Ken Porter. 2015 Mental Health Care System Survey Results. Mood Disorders Society of Canada. Retrieved November 12, 2018 from https://mdsc.ca/research/2015-mental-health-care-system-survey-results/

[30] Robert Reiner. 2008. Integrating a portable biofeedback device into clinical practice for patients with anxiety disorders: results of a pilot study. *Applied Psychophysiology and Biofeedback* 33, 1: 55-61. https://doi.org/10.1007/s10484-007-9046-6

[31] Cecil R. Reynolds, Randy W. Kamphaus, and Kimberly J. Vannest. 2011. Behavior Assessment System for Children (BASC). In *Encyclopedia of Clinical Neuropsychology*. Springer, New York, NY, 366-371. https://doi.org/10.1007/978-0-387-79948-3_1524

[32] Elke A. Schoneveld, Monique Malmberg, Anna Lichtwarck-Aschoff, Geert P. Verheijen, Rutger C. M. E. Engels, and Isabela Granic. 2016. A neurofeedback video game (MindLight) to prevent anxiety in children: A randomized controlled trial. *Computers in Human Behavior* 63: 321-333. https://doi.org/10.1016/j.chb.2016.05.005

[33] David. L. Schwartz and Taylor. Martin. 2006. Distributed learning and mutual adaptation. *Pragmatics & Cognition* 14, 2: 313-332.

[34] Stuart Shanker. 2010. Self-regulation: calm, alert and learning. *Education Canada* 50, 3: 105-138.

[35] Elizabeth A. Shirtcliff, Amber L. Allison, Jeffrey M. Armstrong, Marcia J. Slattery, Ned H. Kalin, and Marilyn J. Essex. 2012. Longitudinal stability and developmental properties of salivary cortisol levels and circadian rhythms from childhood to adolescence. *Developmental Psychobiology* 54, 5: 493-502. https://doi.org/10.1002/dev.20607

[36] Petr Slovák and Geraldine Fitzpatrick. 2015. Teaching and developing social and emotional skills with technology. *ACM Transactions on Computer-Human Interaction* 22, 4: 19. https://doi.org/10.1145/2744195

[37] Naomi J. Steiner, Elizabeth C. Frenette, Kirsten M. Rene, Robert T. Brennan, and Ellen C. Perrin. 2014. Neurofeedback and cognitive attention training for children with attention-deficit hyperactivity disorder in schools. Journal of Developmental & Behavioral Pediatrics 35, 1: 18-27.

[38] Madelon A. Vollebregt, Martine van Dongen-Boomsma, Dorine Slaats-Willemse, and Jan K. Buitelaar. 2014. What future research should bring to help resolving the debate about the efficacy of EEG-neurofeedback in children with ADHD. *Frontiers in Human Neuroscience* 8: 321.