# AssessBlocks: Exploring Toy Block Play Features for Assessing Stress in Young Children after Natural Disasters

XIYUE WANG, Research Institute of Electrical Communication, Tohoku University, Japan
KAZUKI TAKASHIMA, Research Institute of Electrical Communication, Tohoku University, Japan
TOMOAKI ADACHI, Department of Education, Miyagi Gakuin Women's University, Japan
PATRICK FINN, Computational Media Design, University of Calgary, Canada
EHUD SHARLIN, Department of Computer Science, University of Calgary, Canada
YOSHIFUMI KITAMURA, Research Institute of Electrical Communication, Tohoku University, Japan

Natural disasters cause long-lasting mental health problems such as PTSD in children. Following the 2011 Earthquake and Tsunami in Japan, we witnessed a shift of toy block play behavior in young children who suffered from stress after the disaster. The behavior reflected their emotional responses to the traumatic event. In this paper, we explore the feasibility of using data captured from block-play to assess children's stress after a major natural disaster. We prototyped sets of sensor-embedded toy blocks, AssessBlocks, that automate quantitative play data acquisition. During a three-year period, the blocks were dispatched to fifty-two post-disaster children. Within a free play session, we captured block features, a child's playing behavior, and stress evaluated by several methods. The result from our analysis reveal correlations between block play features and stress measurements and show initial promise of using the effectiveness of using AssessBlocks to assess children's stress after a disaster. We provide detailed insights into the potential as well as the challenges of our approach and unique conditions. From these insights we summarize guidelines for future research in automated play assessment systems that support children's mental health.

CCS Concepts: • **Human-centered computing** → *Interactive systems and tools*; **Ubiquitous and mobile computing systems and tools**; • **Applied computing** → **Health care information systems**.

Additional Key Words and Phrases: tangibles for health, toy blocks, stress assessment, children, PTSD, play, well being

**ACM Reference Format:**
Xiyue Wang, Kazuki Takashima, Tomoaki Adachi, Patrick Finn, Ehud Sharlin, and Yoshifumi Kitamura. 2020. AssessBlocks: Exploring Toy Block Play Features for Assessing Stress in Young Children after Natural Disasters. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 1, Article 30 (March 2020), 29 pages. https://doi.org/10.1145/3381016

## 1 INTRODUCTION

Natural disasters are occurring frequently [14, 68], and take a terrible toll: disasters such as the 2004 Indian Ocean Earthquake and Tsunami (228,000 casualties), 2008 Sichuan Earthquake (88,287 casualties), 2010 Haiti Earthquake

Authors' addresses: Xiyue Wang, xwang@riec.tohoku.ac.jp, Research Institute of Electrical Communication, Tohoku University, Sendai, Japan; Kazuki Takashima, takashima@riec.tohoku.ac.jp, Research Institute of Electrical Communication, Tohoku University, Sendai, Japan; Tomoaki Adachi, Department of Education, Miyagi Gakuin Women's University, Sendai, Japan, adachi@miyagi.email.ne.jp; Patrick Finn, Computational Media Design, University of Calgary, Calgary, Alberta, Canada, pfinn@ucalgary.ca; Ehud Sharlin, Department of Computer Science, University of Calgary, Calgary, Alberta, Canada, ehud@cpsc.ucalgary.ca; Yoshifumi Kitamura, Research Institute of Electrical Communication, Tohoku University, Sendai, Japan, kitamura@riec.tohoku.ac.jp.


2474-9567/2020/3-ART30 $15.00
https://doi.org/10.1145/3381016

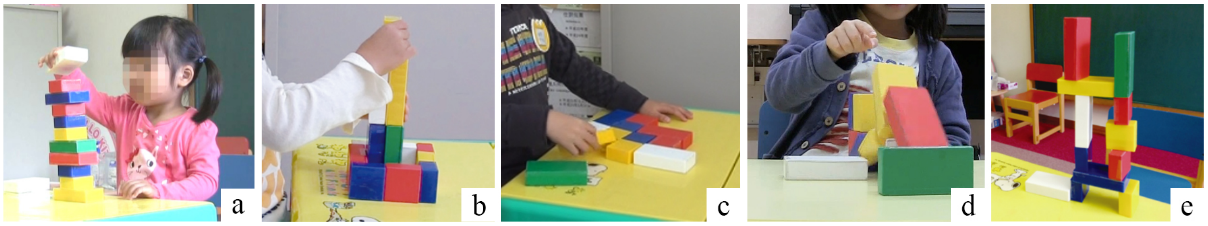



Fig. 1. Post-disaster children's toy blocks play patterns, from left to right: ***(a)*** construction ***(b)*** fence building ***(c)*** playing flat ***(e)*** destruction **(e)** unstable structures

(222,570 casualties), and the 2011 Tohoku Earthquake and Tsunami (22,626 casualties) left behind a complex array of problems, many of which are extremely difficult to solve [28]. Some of the most serious long-term challenges for survivors are mental health problems such as Post-Traumatic Stress Disorder, (or PTSD), conditions which can be particularly serious among young children [43]. Children's mental health problems are reported to be accentuated after significant natural disasters [43], and are often not healed by time [8, 35]. Children are also harder to treat than adults when using traditional, talk-based methods [51]. Young children's linguistic expression and cognitive development are not fully-fledged, compared to adults, so the understanding of children's internal psychological state and related mental health issues requires more delicate observation, and often calls for a different approach.

Researchers have developed computer-based interactive tools for addressing complex issues in children's mental health. In areas such as cognitive impairment, autism, and dyslexia interventions with Tangible User Interfaces (TUIs) have proven particularly effective [20, 29, 63]. While this work makes an important contribution to the study and treatment of mental health in children, less has been done on user interface research that could help those suffering from trauma caused by natural disasters. Given the number of people impacted by these events, and that such events are increasingly common, we propose a novel method to assess children's stress with a TUI approach. TUIs, especially in their simplest form, physical toy blocks, can potentially support play activities that captures some of the physicality of natural disasters - e.g. the physical destruction of real structures. Physical construction and destruction actions are inherent to playing with toy blocks, providing a non-verbal actions that can be mapped directly to the child's inner responses to the traumatic experience.

This paper presents our initial design and evaluation work towards a long-term goal of developing quantitative play feature-characterizing toy blocks for automatically assessing children's mental health. The long-term research question we pursue is:

> Can sensor-enabled toy blocks assess post-disaster stress in children?

Our research questions began to form after observing new patterns in PTSD-affected children playing with toy blocks in kindergarten after 2011 Tohoku Earthquake and Tsunami (see for example [1, 2]). The new patterns included cycles of building block construction followed by intense destructive actions. Six months later, both PTSD symptoms and destructive behavior seemed to diminish. Based on frequent but anecdotal observations, we wondered if the building, destruction, and rebuilding process reflected children's mental and psychological states and might be helping them come to terms with the destruction they witnessed. We consider whether the physicality of the blocks, and the freedom to create structures became a simple medium capturing the children's stress and allowing them to express their anxiety and fears. These observations led us to investigate the connection between children's toy block play patterns and mental state, especially in post-disaster stress. If block-play contained relevant information, perhaps we could automate the block play assessment approach for children's mental healthcare. Automated, computer-based assessment promises to reduce the need for professional assessors' time, improving access by lowering training requirements for assessors, eliminating some forms of bias, and improving the reliability of testing (as demonstrated in [31, 60]). Children's post-disaster automated stress assessment helps extend the quality, affordability, and access to critical health care necessary for many children in a world facing increased exposure to natural disasters.

To realize this vision, we designed and prototyped sturdy, simple, automated blocks with IMU capable of capturing basic play actions (see an example of play and sensor raw data in Figure 2). From 2013 to 2015, two years after the 2011 megathrust in the Tohoku region of Japan, the blocks were deployed in a set of studies with 52 pre-school children, aged 2.7 to 6.9 (see snapshots of studies in Figure 1). Among the participants, 15 (aged 5.9 – 6.5) experienced the most damaging impacts of the tsunami firsthand, when they were between the ages of 3.0 and 4.0. The unprecedented and devastating tsunami placed this group at a higher risk than the other participants for stress-related illness. We sampled approximately 20 minutes of play activity with our blocks, and manually evaluated each child's behavior during the session, noting areas such as concentration or for children who felt lost and required support. We measured each child's stress before and after play, using bio-marker sAA (Salivary Alpha-amylase Activitity), and evaluation form OSBD (Observation Scale of Behavioral Distress) and VAS (Visual Analogue Scale of Anxiety).

Our analysis showed that some of our block features, play behavior evaluations, as well as traumatic experience, related to children's stress measurements. Our findings indicate that the block-play features approach is promising for automatically predicting stress in children.

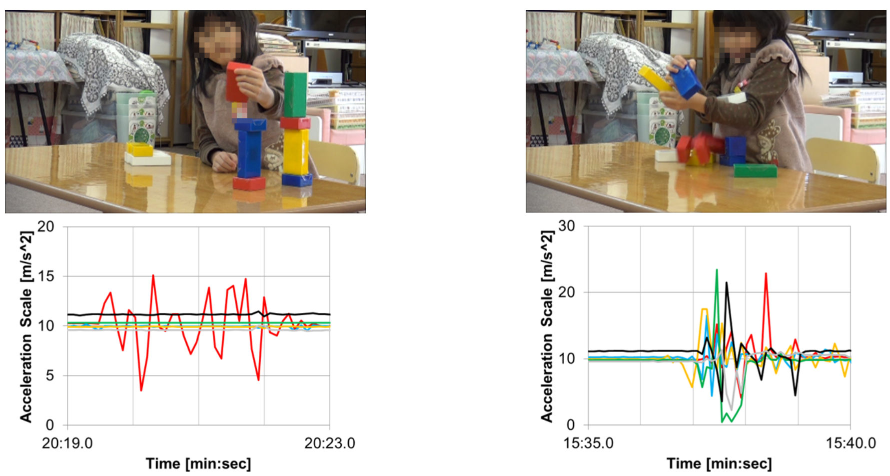



Fig. 2. Children's play and snapshots of data from IMU embedded in our blocks

Our contributions are summarized as follows.

- The design of AssessBlocks, a computer-augmented toy using sensors-embedded toy blocks that document children's block-play features;
- A protocol and procedure for using AssessBlocks to study block play features, behavior evaluations, and children's stress measurements;
- Bivariate analysis of the data we collected with AssessBlocks, revealing correlations between block features, play behavior, trauma experience and stress measurements;
- Discussion of the potential and limitations of AssessBlocks and our toy block assessment approach; a roadmap for further iterations and future research.

## 2 RELATED WORK

Our research is directly motivated by the record-setting Great East Japan Earthquake and its' psychological and social impact on children. The work is built on pediatric, psychological, and social studies of children's mental health after experiencing large-scale traumatic events, HCI research on playful and interactive user interfaces for assisting, treating, and assessing the health of children, and on knowledge of children's activity sensing and characterizing techniques.

### 2.1 Natural Disasters and Their Impact on Children

On March 11, 2011, a record-setting earthquake and tsunami hit the Tohoku (northeast) region of Japan. The Great-East Japan Earthquake is the fourth largest in modern record. At 14:46 JST, a magnitude 9 to 9.1 earthquake struck the coast. Tsunami waves of up to 40.5 meters followed the quake and gave people less than an hour to evacuate.

Japanese buildings are designed to be resilient during earthquakes, and citizens are well-trained for emergency evacuation; however, the scale of this event was a shock to the system. At least 22,626 people lost their lives. Witnesses confirm that children, even those in pre-planned shelters saw people die [32]. The impact of the disaster continues: the children affected by the natural disaster are now teenagers many of whom still suffer from latent PTSD and its long-term symptoms, often exacerbated by repeated exposure to elements of the trauma since major earthquakes are not uncommon in the region.

Surviving a major natural disaster can cause stress and illness in most people but is particularly hard on children [8]. After the 2008 Sichuan earthquake, PTSD and depression rates among children aged 8 to 16 years were 12.4% and 13.9% respectively when measured 15 months after the event [35]. Over 50% of school-aged children in Haiti were reported to have severe post-traumatic stress one year after the 2010 earthquake [8]. Becker's research on trauma in children who experienced the Asian Tsunami on December 26, 2004 showed a tendency towards regressive behaviors [7]. In the study, children between 4 and 7 exhibited typical regressive behaviors such as clinging, bedwetting, fearfulness, sleep disorders, and elevated reactions to stimuli.

Recently, more research attention is focused on needs that arise after survival requirements for food, clothing and shelter are met. In particular, mental health care is receiving increased focus [7, 43, 57, 65, 74]. Coping with disaster-induced mental health conditions is not easy, but timely access to appropriate mental health care is crucial to reduce the risk of developing PTSD [35]. Unfortunately, mental health interventions are usually short-term, and even these are hard to provide and hard to access. They require considerable time, professional expertise, and resources [57]. Many factors further complicate practitioners' ability to target and monitor the appropriate mental health care for children. Natural disasters cause simultaneous systemic shocks, and research shows that cumulative and conjoined traumatic events can amplify behavioral problems [15, 47]. After the 2010 earthquake in Haiti, Blanc relocated children to a center where psychosocial supports were in place. Yet, even with a systematic intervention Blanc's team were unable to show significant reductions in either PTSD or depression when compared to a control group [8]. Pynoos's analysis of PTSD in children following the 1988 Armenian earthquake suggests that girls sustained more serious and long-lasting suffering than boys [52]. While our work focused on helping children who survived the Tohoku earthquake and tsunami, its overarching goal is to promote research that can help all children who experience trauma following natural disasters.

### 2.2 Playful, Interactive Health Assessment and Treatment

Playful or play-based therapies are a well-established means for treating mental health. Creative play approaches such as Sand-Play and Painting Therapy are commonly used to treat chronic stress and PTSD [5, 62]. Block play has shown therapeutic results for social withdrawal and ADHD in children [36, 53]. Pullman has stated that with maturation, young children transition from transporting blocks to stacking them, and then move to three-dimensional composition [50]. As a result, blocks have been used in three-year old children's cognitive development checkups in Japan [40]. Traditionally, play therapy and assessment is conducted by an on-site therapist using observation, followed by question-and-answer interviews with the child, and sometimes involving the analysis of video recordings of their gameplay. These approaches are effective, but are profoundly time-consuming, require advanced therapeutic or psychological assessor expertise, and are often incapable of capturing nuanced differences in play actions.

Computer-assisted health assessment and therapy is increasingly common. Automated or computer-aided assessment can potentially reduce on-site professional time required, and reduce training requirements for caregivers or assessors. For example, "Cognitive Cubes" [60] proposes a method to measure spatial cognitive ability using a tangible interface called "Active- Cube" [41]. The Active-Cube researchers assessed construction ability and dementia by automatically presenting target shapes, then supporting the shape-building process using a 3D construction TUI, and automatically analyzing participants' shape similarity outcomes over time. Intarasirisawat et al. show that touch and motion features collected from three popular mobile games, Tetris, Fruit Ninja and Candy Crush, have potential to be used as proxies for conventional cognitive assessment of such elements as attention and memory[31]. An overview of PTSD diagnoses and treatment suggest that computational technologies can support information gathering and provide more objective PTSD assessments when compared to traditional paper and talk-based methods [44] Exposure therapy using computer simulations shows promising results for prevention and therapy in trauma-related disorders. Botella and Rizzo show the potential of adaptive displays and Virtual Reality Games (VRG) when treating combat-related PTSD [10, 54]. Several systems where caregivers manually track patients have shown that data-driven approaches improve quality of life for those suffering from PTSD and depression [6, 75].

The potential use of TUIs in children's healthcare is being explored in a growing number of projects. In one example, Fan et al. showed that working with tangible letters helped children with dyslexia learn to read and spell [20]. Westeyn et al. created augmented toys called, "Child'sPlay," using Inertial Measurement Units (IMU) and other sensors to support automated recording, recognition, and quantification children's play behaviors for subsequent analysis [73]. While adults use language and various abstractions and representations as their primary means of communicating with the world, TUIs create a unique space for children to express themselves since they are "easier to learn and use", as well as "draw upon physical affordances" and "support cognition through physical representation and manipulation" [25].

Blocks are the most widely accessible play object in early childhood classrooms [12, 50],and a popular form for creating playful interactions among children. Various TUIs were designed to assess and treat children using gameplay with automated blocks. For example, Vonach et al designed "MediCubes" that measure children's physiological parameters during play [70]. Jacoby proposed PlayCubes that assessed children's construction ability using a TUI [33]. StackBlock is a block-shaped interface that detects flexible stacking by embedding a matrix of infrared LEDs and phototransistors [3]. Our approach builds on these past projects and works towards providing young children at-risk of mental health problems after natural disasters a non-verbal TUI-based medium that would allow them to relate to and directly communicate physical elements of their traumatic experience.

## 2.3 Activity Sensing and Detection Techniques

Nowadays, one of the mainstream techniques used in action detection in children is image processing with cameras. Wang, Liu and Yang have presented various techniques for children's activity analysis using distributed cameras and Machine Learning algorithms [45, 71, 76]. While generally effective, there are some common difficulties when retrieving activity information using camera-based methods. First, in the preparation stage, the location of the cameras needs to be well-designed to establish camera views that capture quality information.

Second, during data collection, children's actions are highly flexible and unpredictable, which makes occlusions by objects in the space and on children's bodies difficult to avoid.

Another trend in activity-detection is embedded sensors inside tangibles using Machine Learning methods to model data acquired by the sensors. "Child'sPlay" by Westeyn et al. uses a SVM (support vector machine) to enable the automatic recording, recognition, and quantification of play behaviors in children [73]. Hosoi et al. created IMU-embedded toy blocks, and modeled the raw acceleration data into actions with SVM, to recognize and assess building processes during play with toy blocks [30].

A common problem among all the above Machine Learning-based action-characterization methods happens after data acquisition. It takes considerable time to annotate and label the raw data [73, 76], and the acquired data is often imbalanced since it is extremely hard to ask children to perform the certain tasks [45, 71, 73, 76]. The result is an approach that does not generalize well among all children.

To avoid the above problems, we use sensor-embedded blocks with a state-machine algorithm to characterize different actions. The are several benefits to this approach. Technically, the time and expertise needed for site-specific setup are low, individual blocks are easily identified, and the blocks are durable. The state machine structure frees us from acquiring and labeling the balanced actions. Using simple, durable structures allowed our field studies to progress without excessive preparation and interruptions.

## 3 ASSESSBLOCK DESIGN AND IMPLEMENTATION

In order to enable the investigation of a possible connection between children's toy block play pattern and mental states we used an approach to capture data directly from the play activity as unfolded. We needed to design toy blocks that would retain the familiarity of commonly used wooden toy blocks, but at the same time provide intrinsic tracking of movements and information on the overall playing activity state. The goal was to design a set of simple, sturdy toy blocks that could be quickly deployed in real-world settings, such as kindergartens and used by caregivers, without the need to set up trackers or cameras. We propose a set of specific design guidelines focused on three aspects: appearance, tactile properties, and data acquisition.

### 3.1 Design Criteria

**Appearance:** the blocks must retain the size and appearance of traditional wooden building blocks to preserve the familiarity of the block play experience. Design should follow the traditional toy blocks' color scheme: using primary colors (red, blue and yellow) first, then secondary colors (green, orange, and violet). Blocks also need to be hollow and large enough to embed sensors. Embedding sensors must be completely hidden, including lights and sounds, to avoid causing children to become curious about the inside of the blocks.

**Tactile properties:** the blocks need to be affordable and sturdy enough to endure shaking, dropping and throwing. To preserve the tactility of wooden blocks, the build material should not only look like, but also provide sensory properties, such as hardness, sensory warmth, and thermal behavior of wooden blocks as characterized by Sadoh el al. [55]. The interactive blocks should also match the weight of well-designed commercial products.

**Data acquisition:** the system needs to allow quick and simple deployment in daycares, clinics, or evacuation centers. The sensor embedded in the blocks needs to unobtrusively and robustly capture and transmit human activity data. A good battery charge needs to be maintained during each play session. Even it is not absolutely necessary, we suggest real-time data transmission and monitoring during play sessions. Since the available number of children for study is limited, ensuing optimal sensor operation, and accurate data gathering is crucial.

Our design goals and operational constraints led us to a relatively simple approach: we implemented a set of sturdy Bluetooth IMU-embedded toy blocks (Figure 3), retaining the appearance of familiar wooden toy blocks, allowing for real-time capture of basic play behaviors.

Fig. 3. Toy blocks with the embedded Bluetooth-IMU-sensors

## 3.2 Physical Specification

Our block prototypes are made with PVC Foam Board providing a warm, hard tactile feel. Each wall of a block is sturdily glued. We prototyped two basic shapes: a big block, measuring $100mm \times 50mm \times 25mm$ and $90g$; and, a small block measuring $50mm \times 50mm \times 25mm$, and weighing $45g$. We used paper clay filling to achieve traditional wooden block's weight. The dimensions and the mass (including internal sensor) followed Nichigan Original's Wooden Tsumiki [49], one of the widely available toy block sets on the Japanese market.

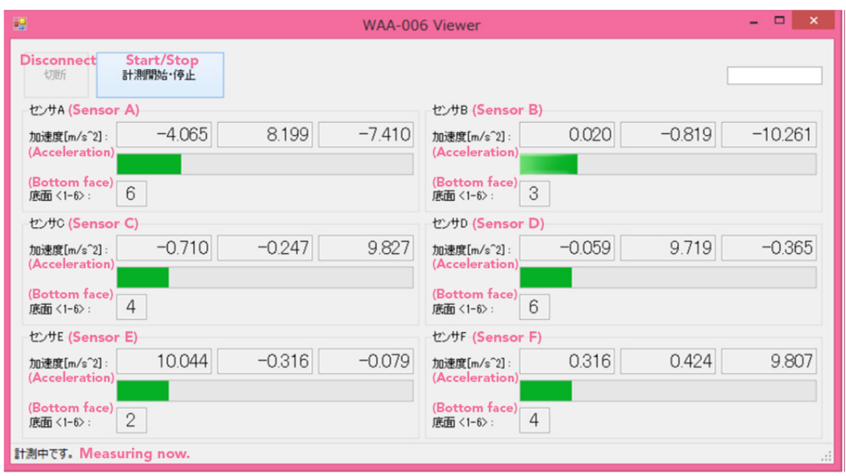



Fig. 4. Software interface on the host computer

## 3.3 Sensor

A wireless IMU sensor is fixed inside each block using Velcro. That, combined with the pressure from the lid, ensures the sensor will not move when shaken or thrown (Figure 3). The wireless IMU sensors (TSND121, ATR-Promotions [4]) hidden in each block have the following specifications:

(a) Move (b) Hold (c) Stand (d) Lay

Fig. 5. Block play actions characterized from pilot study

- Triple axis accelerometer and gyroscope with 3 working axes (X, Y, Z) providing a maximum sampling rate of 1000 Hz;
- Triple axis magnetometers with 3 working axes (X, Y, Z) and maximum sampling rate of 100Hz;
- Maximum acceleration detection range per axis of $\pm 16G$
- Binary format data output;
- Bluetooth communication;
- Built-in battery.

The raw sensor data included x, y, z-axis accelerometer and gyroscope values, which were sent in real-time to a host computer via Bluetooth using a 50Hz frequency, providing data transmission as well as monitoring and inspection capabilities during play sessions. The interface used for data collection is shown in Fig 4.

### 3.4 Block Play Features

We then calculated the numerical activity data from the raw data. To better understand the fundamental play activities that characterize behavior, we conducted a pilot study and observed free block play of 30 kindergarten children. From this study, we derived the following 7 fundamental activity features representing how active and constructive a play session is:

**Time:** total time between the start and stop of the program;

**HoldTime:** total time at least one block is held in hand but not moving. An example of holding a block can be found in Figure 5b;

**MoveTime:** total time at least one of the blocks is moving. An example of moving a block can be found in Figure 5a;

**Movement:** A sum of the magnitude of all three-axis acceleration values within a play session. This value is not equal to velocity since the sensor introduces accumulated error over time, however, it provides data indicating speed variations which are sufficient for comparison between subjects.

We further define two states during which a block is placed. When the largest face of a block contacts the ground, we call it “laying” (see Figure 5d), while “standing” refers to the state when any other face contacts the ground (see Figure 5c).

**StandTime:** the total time when a block is in a “standing” state;

**StandCount:** the number of events when the placing is classified as “standing";

**LayCount:** the number of events when the placing is classified as "laying".

### 3.5 Preprocessing and Play Features Detection

As illustrated in Figure 5, we implemented a threshold-based state machine algorithm to extract total counts and total time for the actions described above.

We first processed raw data to extract difference (Diff) and "Bottom Side". We applied a moving-average filter to the raw accelerometer data using the unweighted mean of 5 data points to filter out high-frequency background noises. We then extracted Diff by (1) applying a sliding window of 20 data points with 50% overlap with the previous window; and, (2) computing Diff as the differences of the average of two adjacent sliding windows. This approach has proven effective as a motion data processing methods [23, 39]. We pre-assigned an ID to each face of a block and used the term Bottom Side to refer to the side facing the ground when a block is placed. Since the raw acceleration data includes gravity, determining the axis gravity was pointing toward allowed us to identify the ID of the Bottom Side for each placement. We then extracted the numerical counts and time of each action using the structure shown in Fig 6.

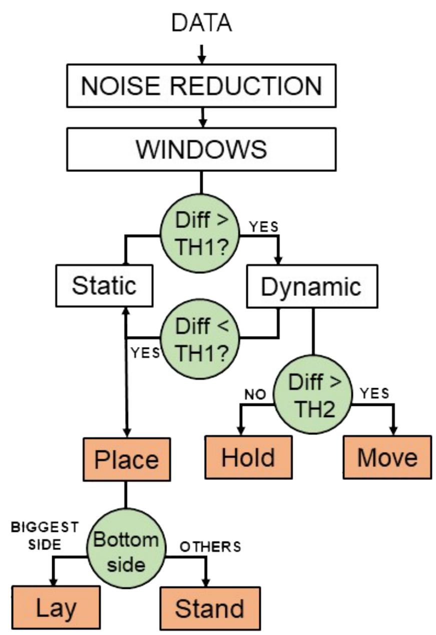



Fig. 6. A threshold-based state-machine structure that recognizes different actions

We next evaluated the accuracy of each action. Time and Movement were not evaluated since data were captured by the sensors and system directly. For count-related actions (StandCount, LanCount), we performed the action 100 times and recorded the count of detected actions. The accuracy is calculated as follows:

$$accuracy = \frac{count_{detect}}{count_{total}}$$

The accuracy is 98.0% for StandCount and 96.0% for LayCount. For time-related action features, we performed the action 10 times, for 10 seconds each time. We compared seconds detected with the ground truth in each trial to establish the error rate for each trial. Accuracy is calculated as follows ($n = 10$):

$$error\ rate = \left| \frac{time_{detect}}{time_{total}} \right|$$

$$accuracy = 1 - \frac{1}{n} \sum_{i=1}^{n} error\ rate_i$$

The accuracy is 98.4% for StandTime, 80.0% for MoveTime, and 66.2% for HoldTime. Play feature detection achieves a high degree of accuracy in Stand and Lay related features, and satisfactory accuracy in MoveTime detection. Accuracy is notably low for Hold Time, for which we suggest the following reasons: HoldTime is hard to detect precisely using acceleration data, with thresholds between movement and stasis. From our observations, a Hold tends to be classified as static, especially during the latter half of a 10-second hold when the holding arm leans against the table making movement minimal. However, since all participants suffer from this offset and children generally do not hold blocks statically for long periods, we consider these anomalies manageable during initial investigations and will investigate methods for improvement.

## 4 AREA-BASED FIELD-STUDY DESIGN

Based on our research question - can sensor-enabled toy blocks assess post-disaster stress in children? we designed a play study that collected children's block play features (from AssessBlocks), children's play-related behavior (through video-coding), and stress measurements measured (on-site and through video-coding) from each play session.

### 4.1 Stress Measurements

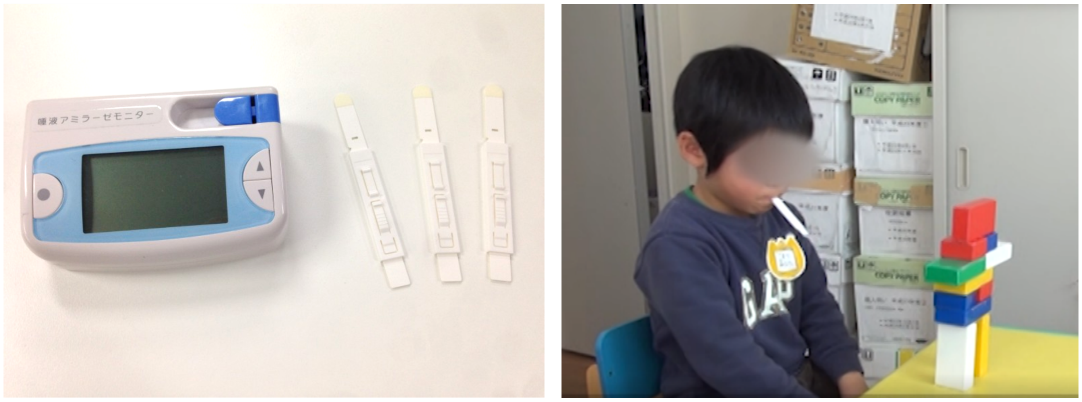
Currently, the best method for collecting stress measurements is self-reporting [16, 38, 66]. However, this method is ineffective with young children who are in the early stages of cognitive and communicative development. As a result, evaluations by professionals and caregivers, combined with bio-markers that can be captured and measured by instruments are often used together to indirectly capture stress in children [22, 24, 61, 67].

In our studies, we captured data using 3 established measurements and biomarkers related to stress:

- Salivary Alpha-amylase activities (sAA);
- Observation Scale of Behavioral Distress (OSBD);
- Visual Analogue Scale of Anxiety (VAS).

**Salivary Alpha-amylase Activities (sAA)** is recognized as a sensitive, but non-invasive biomarker for stress-induced changes in the body connected to activity in the sympathetic nervous system [56]. Alpha-amylase production in the salivary glands increases in response to psychological and physical stress, and has been shown to be an accurate marker of activity in the autonomic nervous system. This approach is commonly used for PTSD-related assessments [21, 48]. In our study, we measured sAA before and after the experiment by asking participants to hold the measurement paper in their mouths for 10 seconds (see Figure 7).

**Observation Scale of Behavioral Distress (OBSD)** is a scale developed to measure children's behavioral responses to events that impact health and wellbeing. OSBD scores can be correlated with ratings of pain, anxiety, and physiological measures before, during, and after such events, and have been effective for evaluating children [19]. We use OSBD for stress evaluation because of its ability to capture subtle changes in a short time [19] since we cannot subject children to long study periods yet need to capture data efficiently. The OSBD measurement is presented as a form as shown in Figure 8a. It takes 13 measurements to capture severe stress responses such as physical resistance, mild stress responses such as asking for help or support, as well as calm play states. The sum of these measurements indicate behavioral stress in a moment of play. During the experiment, the OSBD was



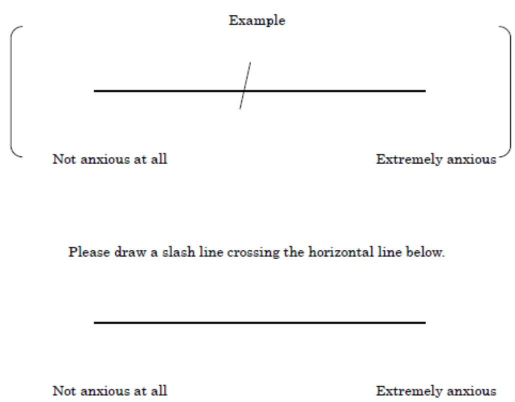
(a) Instrument for measuring sAA

(b) A child is getting sAA measured

Fig. 7. sAA measurments

| | Point |
|---|---|
| 1. Searching for something | 1.5 |
| 2. Crying | 1.5 |
| 3. Crying and Screaming | 4.0 |
| 4. Physical resistance | 4.0 |
| 5. Linguistic resistance | 2.5 |
| 6. Seeking for emotional support | 2.0 |
| 7. Saying painful | 2.5 |
| 8. Aggressively shaking the body | 4.0 |
| 9. Saying scary | 2.5 |
| 10. Body is stiff | 2.5 |
| 11. Nervous behavior | 1.0 |
| 12. Neutral facial expression | 0.5 |
| 13. Smile | 0.0 |

(a) OSBD form




(b) VAS for Anxiety form

Fig. 8. Forms for evaluating children's stress level

measured by an on-site psychologist three times; once at the beginning, once in the middle (approximately 10 minutes later), and then at the end of the play session.

**Visual Analogue Scale of Anxiety (VAS)** measures a characteristic or attitude ranging across a continuum of values. As shown in Figure 8b, VAS captures measurements by asking the observer to draw a vertical line across a horizontal scale indicating the value. The left side of the horizontal scale indicates the minimum value and the right the maximum. VAS is often used in epidemiological and clinical research to measure the intensity or frequency of various symptoms [69]. Its measurement has been used to collect adult's self-evaluation of mood, stress and health across different times of a day [66] as well as parent and staff reports on children's fear, pain and stress [9, 61]. In our study, we asked the child's caregiver to measure the child's anxiety before and after the experiment session.

Our selection of stress measurements was inspired by previous work using these three measurements in combination when evaluating stress in children undergoing medical procedures [61]. The combination of sAA and VAS has also been studied for measuring pain perception in children [22]. Together with sAA, Salivary Cortisol is also commonly used to capture stress bio-markers in children [17, 18, 24, 42], however, we preferred sAA because: (1) it is less invasive, more comfortable, and uses simpler equipment helping us maintain an orderly environment and avoiding a more medical setting which might increase stress in the children [18, 24]; and, (2) Cortisol and sAA are often not correlated [17, 24, 42] while sAA has been shown to be related to other stress measurements such as HR [24], and negative behavior measurements[18]; and, (3) sAA is well-suited for measuring mild to moderate stress responses [24, 67].

Following on previous work [17, 24, 42, 61], we were able to document time-based measurements at before (sAA, OSBD, VAS), during (OSBD only without interrupting the study) and after play (sAA, OSBD, VAS) to collect comprehensive measurements for each session without introducing interruptions.

## 4.2 Play Behavior Measurement

Based on the anecdotal evidence of a change in play behavior in children after the earthquake and tsunami, and on our observations from AssessBlock's pilot study, we designed 4 behavior measurements we speculated would correlate with a child's stress:

- Concentrated Time;
- Lost Time;
- Stacking Time;
- Flat Time.

**Concentrated Time** is calculated by accumulating the time a child is concentrated while playing with blocks, instead of running around, talking, or performing actions that are not block-play related.

**Lost Time** is calculated by accumulating the time a child does not know what to do, seeks help, or looks for encouragement.

**Stacking Time** is the accumulated time that a child is concentrated in stacking and building 3-dimensional structures. Examples are found in Figure 1a and Figure 1b.

**Flat Time** denotes time that a child is merely placing blocks flat down, with the largest side contacting the table. An example is showing in Figure 1c. This behavior was reported as happening with greater frequency immediately following the 2011 event.

The first two categories, Concentrated Time and Lost Time, are often used when examining behavioral conditions such as ADHD in children [36, 53]. The ability to build is commonly used in cognitive development checkups for three-year-olds in Japan [40]. We could not be certain whether we could use these features to assess stress but the approach seemed promising. We worried our approach might over-complicate the implementation of AssessBlock to detect these features, so to account for this we asked psychologists to video-code when measuring OSBD while watching videos of the play session. If proved to be significant identifiers of stress, it would be possible to use AssessBlock to capture data in the future, but we needed to ensure accuracy in the short term.

## 4.3 Participants and Timeline

Six months after the 2011 Earthquake and Tsunami, we started to contact the kindergartens to gather evidence of altered block play behavior, while the development phase of AssessBlock began. Working with childcare and community workers, we designed an area-based field study in one of the most affected areas - Sendai city, Japan.

Fig. 9. Kindergarten rooms for children's study

After getting Ethics agreements approved by affiliated organizations and formal agreements with the parents of each participant, our experiments took place between September 2013 and November 2015.

The participants we recruited were preschoolers aged 2.1 to 6.9 years old, among whom toy blocks are known to be particularly popular [12, 27, 58]. The participants were recruited from the following three locations.

**Coastal kindergarten.** From Oct 2013 to Feb 2014, the play study was conducted with 15 children (aged 5.9 to 6.5) at a kindergarten located in a coastal town. Participants were 3.0 to 4.0 years old on the day of the disaster,. This kindergarten was the one most damaged by the tsunami in the area. On that day, shortly after evacuating to a hill, the first floor of the school building flooded. While waiting for rescue, children watched the tsunami approaching, carrying debris and washing away almost everything in its path. After the event, the kindergarten was closed for two months. Among children lived in the coastal area, many families lost their houses and jobs, and children and their families had to live in temporary shelters for periods of several months to 5 years. The children from this kindergarten were noticeably nervous, passive, and commonly had difficulty concentrating.

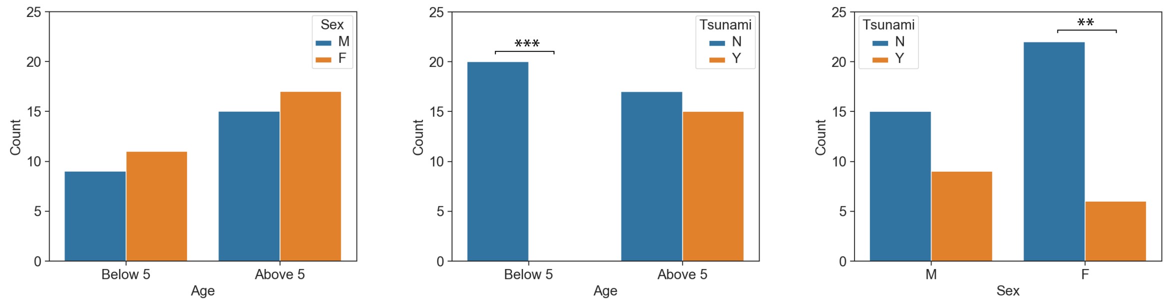
**Inland kindergarten.** From Jan 2014 to April 2014, 17 children (aged 5.1 to 6.9) from an inland kindergarten participated in the study. These children were aged 2.3 to 4.0 on the day of disaster. Children in this group experienced the earthquake, but not the tsunami. Significantly, the building housing their kindergarten was not damaged or interrupted by the 2011 natural disaster.

**Inland children's center.** From Sep 2013 to Nov 2015, 20 children (aged 2.1 to 3.8) were recruited from a children's play center, where children came with their parents for group play and socialization. These children were 0 - 1.3 old at the time of disaster and were out the reach of tsunami, and thus less impacted by the combined effects of the earthquake and tsunami experienced by those from coastal areas.

Whatever the specific experience on that day, all experienced continued aftershocks. As of 16 March 2016 there were 869 aftershocks of over magnitude 5.0; 118 over magnitude 6.0; and, 9 over magnitude 7.0 [34]. The number of aftershocks experienced was shown to be associated with decreased health across Japan [65]. Studies have suggested that the earthquake itself might not as traumatic for the Japanese residents who are resilient to earthquakes, while the unexpected and record-setting tsunami that caused so much death and destruction triggered greater stress and sorrow [32]. Given that direct exposure to traumatic events is a predictor for higher levels of post-traumatic stress, we felt an area-based approach was a good starting point for our research.

### 4.4 Environment, On-site Procedure and Data Collection

All experiments were conducted in the children's familiar environment. Inside the room where the children usually play, a child's desk and chair were prepared and a set of 12 blocks was placed on the desk (see Figure 9). Studies with younger children included a parent, and those in the kindergarten included the students' regular teacher. We kept rooms quiet and well-lit to reduce potential for stress. A pediatric psychologist, and a psychology student were in the room for on-site support and observation. Two HD cameras aimed in different directions captured an audiovisual record of the children playing.




(a) Age (below 5/above 5) of male and female children

(b) Age (below 5/above 5) of children with and without tsunami experience

(c) Gender of children with and without tsunami experience

Fig. 10. Demographic characteristics and trauma experience of the participants. Bars show counts of participants under different demographic characteristics. Statistically significant differences between the counts of different subgroups is denoted by asterisks (** indicates $p < 0.01$, *** indicates $p < 0.001$)

Each child was invited to the desk to play with the AssessBlock toys for a free play session of approximately 20 minutes. Children were free to stop early or continue longer if they wished. Before playing, a parent or a teacher completed consent forms, and filled in a VAS of Anxiety form. The pediatric psychologist completed the OSBD form, and conducted the sAA measurement which includes instructing a child to hold a small paper test strip above the tongue for 10 seconds (Figure 7). The play session then started, and a student research assistant started the AssessBlock program remotely to document the child's block play features. OSBD was evaluated again by the psychologist 10 minutes into the session. After a child stopped, the AssessBlock program was wirelessly stopped, and sAA, OSBD, and VAS were again measured and evaluated.

When the on-site experiments were complete, four more professionals rated each child's OSBD before, during, and after a session by observing the recorded video while video-coding play behaviors for each child. Data from the on-site psychologist, and the four external evaluators were averaged to arrive at a final score.

## 5 RESULTS

In this section, we first present a profile of the data we collected, including demographic factors of the sample, and an examination within each group of data. We then explore the relationships between demographic factors and stress measurements, as well as the association between block-play related features and stress measurements. Spearman's Rank Correlation is used to assess the relationship between a pair of variables since most are either on a ratio or ordinal scale.

### 5.1 Data Profile

*5.1.1 Participants Profile and Trauma Events.* We conducted our experiment with 52 preschoolers (24 male, 28 female), aged 2.1 to 6.9 years old (mean = 5.0, SD = 1.74), from three locations in the Tohoku region of Japan. 15 children (9 male, 6 female) aged 5.9 to 6.6 (mean = 6.4, SD = 0.30) came from the coastal kindergarten. They were aged 3.0 to 4.0 (mean = 3.54, SD = 0.32) when the 2011 events took place. 17 children (6 male, 11 female) aged 5.1 to 6.9 (mean = 6.3 SD = 0.52) were from the inland kindergarten inside the city. They were aged 2.3 to 4.0 (mean = 3.36, SD = 0.52) at the time. Another 20 children (9 male, 11 female), aged 2.1 to 4.4 (mean = 3.0, SD = 0.73) were from an inland children's center and were aged 0 to 1.3 (mean = 0.51, SD = 0.59) at the time of the event. There is

Table 1. Descriptive profile of stress measurements

| Feature | Average | Standard deviation | Range |
|---|---|---|---|
| sAA Before | 67.09 | 55.93 | 4.00 - 240.00 |
| sAA After | 69.92 | 57.44 | 3.00 - 267.00 |
| sAA Ave | 68.51 | 48.33 | 5.50 - 237.00 |
| sAA Diff | 33.50 | 109.13 | -96.00 - 563.64 |
| OSBD Before | 0.85 | 1.05 | 0.00 - 4.00 |
| OSBD Middle | 0.65 | 0.93 | 0.00 - 4.00 |
| OSBD After | 0.45 | 0.67 | 0.00 - 4.00 |
| OSBD Ave | 0.65 | 0.69 | 0.00 - 2.67 |
| OSBD Diff | -0.48 | 1.19 | -3.50 - 4.00 |
| VAS Before | 4.32 | 2.83 | 0.00 - 9.60 |
| VAS After | 2.10 | 2.24 | 0.00 - 8.20 |
| VAS Ave | 3.21 | 2.13 | 0.30 - 8.60 |
| VAS Diff | -2.22 | 2.81 | -7.90 - 7.30 |

a differences in age scale since children from the first two locations are all above 5 years old, while those from the third group are all under 5 years of age (Figure 10a).

At the time of the study, 5 to 6 year olds were the youngest pre-school tsunami victims we could recruit, because those that were older at the time of the tragedy were now in elementary school and were no longer used to playing with blocks.

Studies have shown that traumatic events emerge as a significant contributor to PTSD [15, 21, 47, 59]. In this study, we emphasize the tsunami experience as a traumatic event for participants. Those from our first field study location directly experienced the tsunami. On March 11, 2011, these children were rushed out of their classrooms and evacuated to a hill behind their school to find safety. Children from the other two locations might had experienced the 2011 earthquake but were out the reach of the tsunami.

In the sample, the traumatic event - tsunami experience, is associated with age when age is classified into Above 5 and Below 5 years old ($\chi^2(1) = 10.99$, $p < 0.001$). The trauma experience varied significantly among children below 5 years old ($\chi^2(1) = 20.0$, $p < 0.001$) (see Figure 10b) due to none of those under 5 having witnessed the tsunami.

Among 24 boys (9 with tsunami experience) and 28 girls (6 with tsunami experience), we did not observe a significant association between tsunami experience and gender; however, the traumatic experience varied significantly within the girls' cohort ($\chi^2(1) = 9.14$, $p = 0.002$ ) (see Figure 10c)

*5.1.2 Stress Measurements.* The stress biomarker sAA is measured at the beginning and end of each session, resulting in two features - sAA Before and sAA After. VAS from a caregiver or a kindergarten teacher is collected at the beginning and end, resulting in two features - VAS Before and VAS After. OSBD is evaluated at the beginning, middle, and end, both on-site and using the video coding, resulting in three features OSBD Before, OSBD Middle and OSBD After.

We then average the features of each stress measurements, to obtain a stress indicator for the entire session - sAA Ave, OSBD Ave, VAS Ave (Ave: an abbreviation for Average). We also calculated the changes of each measurement over the play session for each individual. For OSBD and VAS, we define the differences as after-values minus before-values, resulting in OSBD Diff and VAS Diff (Diff: an abbreviation for Differences). For sAA,

Table 2. Correlations among stress measurements

| | sAA Before | sAA After | sAA Ave | sAA Diff | OSBD Before | OSBD Middle | OSBD After | OSBD Ave | OSBD Diff | VAS Before | VAS After | VAS Ave | VAS Diff |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| sAA After | **0.520**$^{***}$ | | | | | | | | | | | | |
| sAA Ave | **0.824**$^{***}$ | **0.889**$^{***}$ | | | | | | | | | | | |
| sAA Diff | **-0.348**$^{*}$ | **0.536**$^{***}$ | .157 | | | | | | | | | | |
| OSBD Before | 0.026 | 0.009 | 0.051 | 0.085 | | | | | | | | | |
| OSBD Middle | -0.263 | -0.131 | -0.216 | 0.095 | **0.527**$^{***}$ | | | | | | | | |
| OSBD After | **-0.348**$^{*}$ | 0.010 | -0.148 | **0.309**$^{*}$ | **0.457**$^{***}$ | **0.819**$^{***}$ | | | | | | | |
| OSBD Ave | -0.174 | -0.057 | -0.114 | 0.160 | **0.784**$^{***}$ | **0.889**$^{***}$ | **0.813**$^{***}$ | | | | | | |
| OSBD Diff | **-0.274**$^{*}$ | -0.097 | -0.208 | 0.027 | **-0.700**$^{***}$ | -0.207 | 0.107 | **-0.396**$^{**}$ | | | | | |
| VAS Before | -0.205 | -0.131 | -0.139 | -.009 | **0.314**$^{*}$ | 0.167 | 0.216 | 0.229 | -0.136 | | | | |
| VAS After | -0.107 | 0.166 | 0.078 | 0.248 | 0.150 | **0.380**$^{**}$ | **0.396**$^{**}$ | **0.278**$^{*}$ | -0.001 | **0.479**$^{***}$ | | | |
| VAS Ave | -0.199 | -0.019 | -0.076 | 0.109 | **0.275**$^{*}$ | **0.346**$^{*}$ | **0.368**$^{**}$ | **0.333**$^{*}$ | -0.100 | **0.878**$^{***}$ | **0.802**$^{***}$ | | |
| VAS Diff | 0.183 | 0.243 | 0.212 | 0.140 | -0.143 | 0.098 | 0.036 | -0.001 | 0.059 | **-0.756**$^{***}$ | 0.057 | **-0.475**$^{***}$ | |

*Note:* $^{*}$: $p$ <0.05, $^{**}$: $p$<0.01, $^{***}$: $p$<0.001, bold values indicate significant correlations.

we calculated the percentage change by taking the difference between post- and pre-play and dividing by the post value, resulting in sAA Diff. In general, a positive Diff indicates an increased stress measurement, while a negative Diff shows a decrease. The average, standard deviation and range values of the 13 stress measurements across 52 participants are presented in Table 1.

To briefly investigate the association between three stress measurements, which includes 4 sAA features, 5 OSBD features, and 4 VAS features, Spearman's rank correlation coefficients are computed between pairs of variables, as shown in Table 2.

The measurements within each stress category are highly correlated. Before and After measurements are positively correlated in sAA (r = 0.52, p < 0.001), OSBD (r = 0.46, p < 0.001) and VAS (r = 0.48 p <0.001). The results indicate that a child with a relatively high stress measurement before a session tended to have a relatively high stress measurement after a session. Averages of all three stress measurements are positively correlated to

Table 3. Descriptive profile of block and play behavior features. Time feature is documented in *min*, and other features are documented in /*min*

| Feature | Average | Standard deviation | Range |
|---|---|---|---|
| P.Conc[a] | 0.93 | 0.16 | 0.26 - 1 |
| P.Lost | 0.30 | 0.28 | 0 - 1 |
| P.Stack | 0.82 | 0.21 | 0 - 1 |
| P.Flat | 0.35 | 0.27 | 0 - 0.95 |
| B.Time[b] | 19.97 | 3.26 | 8.18 - 26.54 |
| B.MoveTime | 0.79 | 0.20 | 0.30 - 1 |
| B.HoldTime | 0.17 | 0.13 | 0.01 - 0.55 |
| B.StandTime | 0.58 | 0.27 | 0.05 - 1 |
| B.StandCount | 12.87 | 5.51 | 3.88 - 27.92 |
| B.LayCount | 11.59 | 4.23 | 2.41 - 22.96 |
| B.Movement | 23.22 | 8.63 | 6.61 - 48.98 |

[a]P. = Play behavior. Conc = Concentration Time
[b]B. = Block feature

their Before and After measurements, which matches expectation since Ave is a combination of Before and After and these two are shown to be correlated. Diff is negatively correlated to Before value in sAA ($r = -0.35$, $p < 0.05$), OSBD ($r = -0.70$, $p < 0.001$) and VAS ($r = -0.756$, $p < 0.001$), indicating that those with a high stress value before the session reduce more stress during the session. Diff is positively correlated to After value in sAA ($r = 0.54$, $p < 0.001$), indicating that those who have a low sAA value after the session show greater reductions during the session. Meanwhile, Diff is negatively correlated to Ave value in OSBD ($r = -0.40$, $p < 0.01$) and VAS ($r = -0.48$, $p < 0.001$), indicating that those with greater OSBD and VAS reductions in the session tend to have a low average value.

We then notice that OSBD and VAS are positively correlated in Before ($r = 0.31$, $p < 0.05$), After ($r = 0.40$, $p < 0.001$), and Ave ($r = 0.33$, $p < 0.05$), while similar correlations are not found in the other two pairs - OSBD and sAA, VAS and sAA. None of VAS measurement is correlated to sAA, while some OSBD measurements are correlated to sAA indirectly: OSBD After is correlated to sAA Before ($r = -0.35$, $p < 0.05$) and sAA Diff ($r = 0.30$, $p < 0.05$), and OSBD Diff is correlated to sAA Before ($r = -0.28$, $p < 0.05$).

Fundamentally, sAA measures physiological and psychological stress while OSBD and VAS measures observable behavioral stress. Based on the above observations, two behavioral stress measurements OSBD and VAS agree in Before, After, and Ave, while the same pattern is not observed between stress biomarker sAA and behavioral stress measurements.

OSBD is evaluated in a more objective and unbiased manner in comparison with to VAS. To simplify the dimension of our targets, in the following analysis we use OSBD measurements to indicate behavioral stress, and we use sAA to indicate physiological and psychological, or inner stress. To further reduce dimensions, we also omit OSBD Mid value since it is highly correlated to OSBD Before ($r = 0.53$, $p < 0.001$), After ($r = 0.82$, $p < 0.001$), Ave ($r = 0.89$, $p < 0.001$), and is included in the calculation of Ave. A total of 8 measurements, including stress biomarker, sAA Before, After, Ave, Diff, and behavioral stress, OSBD Before, After, Ave, Diff, are used in the following analysis.

Table 4. Correlations among play behavior, block features, and between them

| | P. Conc | P. Lost | P. Stack | P. Flat | B. Time | B. MoveTime | B. HoldTime | B. StandTime | B. StandCount | B. LayCount | B. Movement |
|---|---|---|---|---|---|---|---|---|---|---|---|
| P. Lost | 0.022 | | | | | | | | | | |
| P. Stack | 0.248 | 0.156 | | | | | | | | | |
| P. Flat | 0.169 | 0.147 | **-0.642**$^{***}$ | | | | | | | | |
| B. Time | 0.065 | 0.024 | 0.247 | -0.197 | | | | | | | |
| B. MoveTime | -0.185 | -0.116 | -0.197 | 0.109 | -0.173 | | | | | | |
| B. HoldTime | -0.046 | -0.168 | 0.221 | -0.248 | **0.348**$^{*}$ | 0.076 | | | | | |
| B. StandTime | 0.086 | 0.204 | 0.142 | -0.107 | -0.140 | **-0.541**$^{***}$ | -0.241 | | | | |
| B. StandCount | 0.068 | **0.298**$^{*}$ | **0.423**$^{**}$ | **-0.277**$^{*}$ | 0.113 | **-0.389**$^{**}$ | **0.368**$^{**}$ | **0.481**$^{***}$ | | | |
| B. LayCount | 0.077 | 0.107 | **0.450**$^{***}$ | **-0.402**$^{**}$ | 0.255 | -0.157 | **0.701**$^{***}$ | 0.048 | **0.794**$^{***}$ | | |
| B. Movement | 0.071 | 0.190 | **0.557**$^{***}$ | -0.270 | **0.314**$^{*}$ | -0.129 | **0.649**$^{***}$ | 0.061 | **0.654**$^{***}$ | **0.745**$^{***}$ | |

*Note:* $^{*}$: $p$ <0.05, $^{**}$: $p$<0.01, $^{***}$: $p$<0.001, bold values indicate significant correlations.

*5.1.3 Block Features and Video-Coded Play Behavior.* Block features, in time and count, and play behavior, in time, are computed across a play session. Since the total length of a session differs between individuals, we further divide each feature, excepting the total time, by the total time in minutes, to obtain features per minute. The average, standard deviation and range values of the 7 block play features and 4 play behaviors across 52 participants are presented in Table 3.

The association among play behaviors, among block play features, and between them are investigated using Spearman's rank correlation. Their bivariate coefficients are shown in Table 4.

The 4 play behaviors are relatively independent, except that Stack is negatively correlated with Flat ($r = -0.64$, $p < 0.001$). This indicates that a child who dedicates more time to stacking spend relatively less time playing flat and vice versa.

In 7 block features, several correlations were found. First, StandCount is positively correlated to LayCount ($r = 0.79$, $p < 0.001$). Contrary to correlations of play behaviors Stack and Flat, counts of standing and laying blocks are highly correlated in the same direction. StandCount is positively correlated to Movement ($r = 0.65$, $p < 0.001$) while negatively correlated to MoveTime ($r = -0.39$, $p < 0.01$). This indicates that children who stand blocks more also moving them faster though not more frequently. StandCount is also positively correlated to StandTime ($r = 0.48$, $p < 0.001$) and HoldTime ($r = 0.37$, $p < 0.01$). HoldTime is positively correlated to LayCount (r = 0.70, p <

Table 5. Correlations between demographic factor, trauma event, and stress measurements

| | sAA Before | sAA After | sAA Ave | sAA Diff | OSBD Before | OSBD After | OSBD Ave | OSBD Diff |
|---|---|---|---|---|---|---|---|---|
| Tsunami Exp | **-0.356**$^{**}$ | -0.235 | **-0.306**$^{*}$ | 0.211 | **0.313**$^{*}$ | 0.190 | 0.209 | -0.104 |
| Age Above5 | -0.259 | -0.257 | -0.252 | -0.016 | -0.057 | 0.119 | -0.031 | 0.119 |
| Age | -0.216 | -0.227 | -0.244 | 0.004 | -0.080 | -0.028 | -0.127 | 0.087 |

*Note:* $^{*}$: $p$ <0.05, $^{**}$: $p$<0.01, $^{***}$: $p$<0.001, bold values indicate significant correlations.
Tsunami Exp = tsunamp experience (0 = No, 1 = Yes)

0.001), Movement (r = 0.65, p < 0.001), and Time (r = 0.35, p < 0.05). Additionally, LayCount and Movement (r = 0.75, p < 0.001), StandTime and MoveTime (r = -0.54, p < 0.001), and Time and Movement (r = 0.31, p < 0.05) are shown to be correlated.

Between block features and play behavior features, we observe that among 4 play features, Stack and Flat are correlated to some block features. The play behavior Stack is positively correlated to LayCount (r=0.45, p<0.001) and StandCount (r = 0.42, p < 0.01), while Flat is negatively correlated to both LayCount (r = -0.40, p < 0.01) and StandCount (r = -0.28, p < 0.05). Movement is positively correlated to Stack (r = 0.56, p < 0.001), indicating that children who moves faster tend to do more stacking. Besides Stack and Flat, StandCount is found positively correlated with Lost (r = 0.30, p < 0.05).

In general, many block features are correlated. Those showing the most correlation with others include StandCount and HoldTime. We also found that block features are commonly related to the play behaviors stacking and playing flat, but not with concentrated time.

## 5.2 Correlations with Stress Measurements

*5.2.1 Demographic Factors, Trauma Events and Stress.* Previous studies show that socio-demographic factors, such as age and gender operate as significant mediators of PTSD [11, 52, 64], and trauma events emerge as a significant contributor to PTSD [15, 21, 47, 59]. Therefore, we investigated whether or not correlations exist between socio-demographic factors and stress measurements.

We did not observe any significant differences in stress measurements between male and female children using a one-way ANOVA test. From our participant profile, we note that in the data set, those with tsunami experience are concentrated among children over 5 years of age. Thus, we speculate that if a stress measurement is found to correlate with tsunami experience, there is a chance that the correlation is influenced by to the stress measurement's correlation with age, notably, with differences in stress below and above the age of 5. As a result, we added an additional ordinal demographic variable, Age Above 5, to account for this condition. To investigate the relationship between age, tsunami experience, and stress measurements, Spearman's rank correlation coefficients are computed between Tsunami Exp, Age, Age Above 5, and the 8 stress measurements.

As shown in Table 5, Tsunami Exp is negatively correlated to sAA Before (r = -0.36, p < 0.01) and sAA Ave (r = -0.306, p < 0.05), and positively correlated to OSBD Before (r = 0.313, p < 0.05). No correlations are found between age and stress measurements, and between Age Above 5 and stress measurements, indicating that the correlation between traumatic event, tsunami experience, and stress measurement are not due to the correlation between age and stress measurements. Notably, Tsunami Exp contribute differently to sAA and OSBD - where

Table 6. Correlations between demographic factor, trauma event, and block features, play behavior

| | P. Conc | P. Lost | P. Stack | P. Flat | B. Time | B. MoveTime | B. HoldTime | B. StandTime | B. StandCount | B. LayCount | B. Movement |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Tsunami Exp | **0.283**$^{*}$ | **0.350**$^{*}$ | 0.117 | 0.208 | -0.075 | 0.061 | 0.021 | 0.027 | 0.177 | 0.115 | **0.321**$^{*}$ |
| Age | **0.318**$^{*}$ | **0.443**$^{**}$ | **0.410**$^{**}$ | 0.026 | -0.107 | -0.152 | -0.183 | 0.232 | **0.302**$^{*}$ | 0.164 | **0.371**$^{**}$ |

*Note:* $^{**}$: $^{*}$: $p$ <0.05, $^{**}$: $p$<0.01, $^{***}$: $p$<0.001, bold values indicate significant correlations..

correlations indicate that those with the tsunami experience have a lower sAA Before and sAA Ave, and a higher OSBD Before value.

*5.2.2 Block Features, Play Behavior, Demographic Factors and Traumatic Events.* Previous research suggests that block playing behavior differs across age and gender [40, 50], and boys and girls have different block playing behavior [13, 26]. Since demographic factors and traumatic events may work as mediators for stress, we briefly investigate the correlation between block features, play behavior and those factors.

We first test whether there are significant differences in block features and play behaviors between genders using one-way ANOVA. We do not observe significant differences between genders in any play behaviors. Among block features, HoldTime ($F(1,50) = 7.822$, $p < .001$), Movement ($F(1,50) = 7.088$, $p < .01$), and LayCount ($F(1,50) = 4.575$, $p < .05$) are significant different between genders.

As shown in Table 6, among play behaviors, Age is positively correlated to Conc ($r = 0.318$, $p < 0.05$), Lost ($r = 0.443$, $p < 0.01$) and Stack ($r = 0.410$, $p < 0.01$). Among block features, Age is positively correlated to StandCount($r = 0.302$, $p < 0.05$) and Movement ($r = 0.371$, $p < 0.01$). Predictably, StandCount, which indicates the 3-dimensional construction, and Movement, which indicates moving speed, increases with the maturation of a child.

Among play behaviors, Tsunami Exp is positively correlated to Conc ($r = 0.283$, $p < 0.05$), Lost ($r = 0.350$, $p < 0.01$). Among block features, Tsunami Exp is only positively correlated to Movement ($r = 0.321$, $p < 0.05$).

*5.2.3 Block Features, Play Behaviors and Stress.* Spearman's rank-order correlation coefficients between 7 block features, 4 play behavior, and stress measurements are presented in Table 7. Between behavior features and physiological stress measurements, Flat is positively correlated to sAA Diff ($r = 0.28$, $p < 0.05$). Notable correlations exist between Conc and sAA Before ($r = -0.27$ $p = 0.053$), and Flat and sAA Before ($r = -0.26$ $p = 0.06$) though they are not statistically significant since the p-values are slightly greater than 0.05. Interestingly, the negative correlations between Flat and sAA Before, and positive correlation between Flat and sAA Diff indicate that a child who has more time playing without construction tends to have a lower sAA starting value, and greater increase of sAA or less of a reduction after the session.

Among block features, sAA is negatively correlated to StandTime ($r = -0.303$, $p < 0.05$), and StandCount ($r = -0.345$, $p < 0.05$). While StandCount and StandTime both point to active construction, above finding indicates that sAA After are lower among those who construct; however, the same observation is not found in other sAA measurements.

For the behavioral stress measurement OSBD, there were on correlations with play behaviors. OSBD After is negatively correlated to the block features Time ($r = -0.44$, $p < 0.001$) and MoveTime ($r = -0.28$, $p < 0.001$). OSBD Ave is negatively correlated to Time ($r = -0.32$, $p < 0.05$). Surprisingly, none of the construction-related block features are correlated to the behavioral stress measurement OSBD.

Table 7. Correlations between block features, play behavior, and stress measurements

| | sAA Before | sAA After | sAA Ave | sAA Diff | OSBD Before | OSBD After | OSBD Ave | OSBD Diff |
|---|---|---|---|---|---|---|---|---|
| P. Conc | -0.269! | -0.221 | -0.243 | -0.011 | 0.102 | -0.066 | -0.114 | -0.039 |
| P. Lost | -0.089 | -0.092 | -0.055 | -0.017 | 0.101 | 0.153 | 0.080 | -0.022 |
| P. Stack | 0.078 | -0.133 | -0.003 | -0.248 | 0.009 | -0.222 | -0.229 | 0.003 |
| P. Flat | -0.259! | 0.010 | -0.125 | **0.280*** | 0.002 | 0.168 | 0.110 | 0.071 |
| B. Time | 0.219 | 0.160 | 0.226 | 0.014 | -0.060 | **-0.444***** | **-0.317*** | -0.242 |
| B. MoveTime | -0.095 | 0.107 | 0.014 | 0.118 | 0.068 | **-0.281***** | 0.245 | 0.076 |
| B. HoldTime | 0.094 | -0.056 | -0.017 | -0.073 | -0.095 | -0.208 | -0.251 | 0.089 |
| B. StandTime | -0.082 | **-0.303*** | -0.255 | -0.151 | 0.113 | -0.010 | 0.099 | -0.137 |
| B. StandCount | -0.013 | **-0.345*** | -0.220 | -0.163 | 0.010 | -0.190 | -0.138 | -0.035 |
| B. LayCount | -0.050 | -0.254 | -0.190 | -0.103 | -0.075 | -0.185 | -0.220 | 0.042 |
| B. Movement | 0.076 | -0.132 | -0.039 | -0.087 | -0.066 | -0.235 | -0.247 | -0.021 |

*Note:* *: $p$ <0.05, **: $p$<0.01, ***: $p$<0.001, !: $p$<0.1,bold values indicate significant correlations.

In general, 2 out of 7 block features, and 2 out of 4 play behaviors show correlation with the physiological and psychological stress sAA. Only 2 from block features, show significant correlation with behavioral stress OSBD. Though not optimal, the results indicate the potential for assessing stress, and particularly internal stress as captured by sAA using toy block play features.

## 6 DISCUSSION

Our study explores a novel approach using sensor-embedded blocks to assess children's stress two years after a large-scale earthquake and tsunami. Our findings provide an encouraging refection on the potential of the approach and help us form and propose guidelines for further research into the assessment of children's mental health using sensor-embedded block approaches. Here we discuss the potential and challenges of this work, the lessons learned, and provide guidelines for future exploration of this area of research.

## 6.1 Potentials

*6.1.1 Addressing the long-term RQ: can sensor-enabled toy blocks assess post-disaster stress in children?* This exploratory study and correlation analysis show the potentials of block play features for assessing physiological and behavioral stress in children.

Focusing on block features, we find the 3D construction-related features, StandTime and StandCount, are negatively correlated to sAA After, indicating that fewer standing-related actions are associated with a higher sAA After, and vice versa. In play behavior features, Flat is negatively correlated to sAA Before and positively correlated to sAA Diff, indicating that more "playing flat" is associated with a relatively lower sAA Before value and an increased sAA during the play session. Simultaneously, from our block and play behavior feature correlations, we find Flat is negatively associated with Stack and StandCount. These combined findings indicate that relatively passive play - with more playing on a flat surface, and less 3D construction actions such as standing blocks up, tends to indicate a lower sAA Before, and a higher sAA After. In such cases, children's physiological stress sAA tends to increase.

Our findings also indicate that children who were directly exposed to the tsunami also exhibit lower sAA Before and Ave. In previous studies, Feldman et al. found that children exposed to war who were diagnosed with PTSD exhibited a low-level of stress biomarker sAA, at baseline, following a challenge, and during recovery [21]. Their results though focused on children exposed to a different kind of traumatic event, seem to align with our findings of sAA value among those with direct tsunami exposure, which indicate those who directly experienced the tsunami as a trauma event might be at greater risk for PTSD. Feldman et al. also discovered that children without PTSD employ comfort-seeking strategies while children with PTSD withdraw [21]. Their findings with regard to withdrawal in children with PTSD also seems to align with the "passive play" behavior we observed, and the low sAA Before value associated with "passive play". This result could indicate that "passive play" might be an indicator for children suffering from latent PTSD.

We also observe that the play behavior Concentrated Time is negatively correlated to sAA Before, similar to Flat. Meanwhile, Concentrated Time does not exhibit a positive correlation to sAA Diff as "playing flat" does. Without correlation to Flat, Concentrated Time may not relate to the "passive play" behavior mentioned above. Thus, the length of time a child is concentrated during play may work as a general indicator of physiological and psychological stress, but not as an indicator of "passive play" and the corresponding bio-marker tendencies.

For behavioral stress - OSBD, correlations between its measurements and play behavior features were not found. The only correlations found were between OSBD and computed block features. While most construction-related features do not exhibit a correlation with OSBD, Time and MoveTime are both negatively correlated to OSDB After. Since block features Time and MoveTime are not significantly correlated, their negative correlations with OSBD measurements should be examined independently. We set up our play sessions to run for 20 minutes, but left flexibility for children to choose how long they engaged in order to be as gentle with them as possible. Predictably, some left early and others played longer with total experiment time varing from 8.18 minutes to 26.54 minutes. The negative correlation between OSBD After and Time indicates that those who ended early, probably due to losing interest, tended to have a higher OSBD after play and a higher average OSBD, while the opposite was true for those still immersed in play after 20 minutes. Additionally we noted that those who had longer periods during which at least one block was moving indicated a lower OSBD measure at the end, and vice versa, revealing a negative relationship between active play and a high level of behavioral stress OSBD.

While in the early stages, our exploratory findings suggest a connection between block play actions and stress in children. There seems to be significant evidences to support the possibility of an automated block system that support assessment and maybe even predictions of children's stress in children after natural disasters.

*6.1.2 Stress Measuring Method.* In this early exploration related to children's mental health, and in particular to PTSD related health issues in children, we used three stress measurements. Distinct from questionnaire and self-evaluation method usually used in studies with adults [66], we used biomarker measurement and behavior evaluation by a third party to approximate stress levels in child. We felt that these were appropriate and relatively unobtrusive methods for working with children and believe they induces less stress fluctuations as a result of running the experiment itself.

We found no convergence between three measurements; however, OSBD and VAS seemed to agree in Before and After values, which validates the credibility of our behavioral measurements. We also found a negative correlation between the Before value of physiological stress sAA and Before value of behavioral stress OSBD. With that, it is hard to simply say the best ground truth for stress is OSBD.

In this study, we consider that OSBD, and particularly OSBD After, captures a child's external and observable stress in daily life, since OSBD Before values may be impacted by unavoidable environmental factors such as being asked to step into an experimental setting. We believe sAA is a promising internal measurement, which might be highly relevant for to traumatic experiences leading to an increased risk of PTSD. sAA reflects on the internal nervous system activity and was found to be related to the reactions to trauma experience and latent stress disorder [21]. In our study, sAA correlated to certain behaviors such as "passive play"; however, the dissonance between sAA and OSBD, and the hidden relationships between external behavioral stress and internal physiological stress requires future investigation.

*6.1.3 Mediator and Moderator Effects.* A mediator is a bridge between a pair of variables while a moderator regulates the size and direction of the association between two variables [46]. In PTSD research, demographic factors and exposure to traumatic events are often shown to be mediators, moderators, or both [11, 15, 21, 47, 52, 59, 64].

Our findings show that tsunami exposure correlates to both sAA and OSBD, and to block features and play behavior, indicating that tsunami experience may operate as a mediator for stress. Some block and play behavior features also correlate to age and gender, but age and gender do not exhibit the mediation effect since no correlation with stress biomarker or behavioral stress was found; however, these features may work as moderators influencing the correlation between play and stress differently (*i.e.,* among different ages, or genders). Moreover, other mediator and moderator factors such as trauma history, family social status, parental socio-demographic factors, and other relevant demographic details may exist.

*6.1.4 Block Features and Play Behavior Features.* From our analysis, we find both block features and play behavior features are correlated to stress measurements; however, human-annotated play behaviors did not outperform block features as we expected. Nevertheless, play behavior features did uncover a "passive play" behavior, and indicated that stress may be possibly related to concentration. Thus, play behaviors seem to perform an important role in stress assessment. Extracting these features automatically from AssessBlocks, either by investigating the computational algorithm, or by incorporating new sensors will be necessary to continue this work.

*6.1.5 The Predictive Capability of Features.* Using the same set of data, previous work shows that block features are reliable for predicting age [37], confirming previous studies of differences in block play between age groups [50] and demonstrating the predictive power of automated blocks. With correlations shown between block features and stress measurement, we see potential for predicting stress and other mental health factors using block features. Even though correlations between block features and stress measurements are not strong, we believe the aggregation of several sets of features may provide the required predictive power. Stress might be predictable using block features and play behaviors, mediated and moderated by demographic factors and trauma experience.

A practical obstacle to examining predictions, particularly with machine learning techniques, is the relatively small data size. Sampling psychological measurements to attain the large datasets necessary for constructing advanced machine learning models is challenging by nature. One way to continue to explore the prediction of stress measurement is to maintain focus on simple and interpretable methods such as linear models and simple nonlinear approaches such as Decision Trees. Another possibility is to explore Multitask Learning, where multiple correlated learning tasks can be solved simultaneously [66]. Since our stress measurement targets are highly correlated, particularly on Before, After and Ave, we are able to multitask targets to obtain a model for prediction.

*6.1.6 Block Assessment Approach.* All of the children showed interest in playing with our block prototypes, and none tried to open or break them indicating that our design approach achieved its goals. No child demonstrably rejected our study method indicating a level of comfort with our approach to field studies. These small signs seem to validate our toy block-based data collection and assessment approach as non-invasive and child-friendly which are crucial for both ethical and methodological reasons.

## 6.2 Limitations and Challenges

While our results are promising, more work is needed in order to use AssessBlocks as a tool to support mental health assessment. A number of limitations need to be addressed in further iterations.

*6.2.1 Data Sampling.* In this work, the data size is relatively small, and is unbalanced with regard to exposure to the traumatic event. The traumatic event in our study is particularly important, since Tsunami Experience is significantly correlated to both physiological stress and behavioral stress. In our data set, 15 among 52 had a tsunami experience. As the result of several constraints, all those with tsunami experience came from one location, and all were above the age of 5. This group is the youngest tsunami victim pre-schoolers we could access under the conditions we faced. The coastal kindergarten group was unique, as the only one in Sendai area damaged by the tsunami, and comparable collaborators with similar circumstances would be difficult to find and engage. While we cannot reasonably access comparative data sets, we believe collecting the data from a representative place affected by the natural disaster was crucial. With the data we had, we observed that tsunami experience is significantly associated with stress and age is not. Balanced data, with more children who experienced the tsunami from a wide range of locations would further validate our findings of correlations between tsunami experience and stress, and would help evaluate the predictive capability of our approach. At the same time, given the sensitivity of the topic and those affected, broadening the sampling size presents unique challenges.

*6.2.2 Block Features and Play Behavior Features.* Some block features, such as HoldTime, demonstrated relatively low accuracy (66.2%) which needs to be addressed. We believe this rate exposes a weakness in threshold-based motion extraction. One of the potential solutions could be introducing sensing modalities to the surfaces of the blocks. As previously illustrated, many block parameters are correlated, with significant Spearman's Rank Order coefficients from 0.31 to 0.79. While it is natural that some actions are related, such highly correlated parameters are usually considered detrimental to statistical analysis and could falsely increase the fit if used in linear models. Thus, a feature selection process is needed for stress prediction with our current set of features. In order to capture play events from different angles, we can explore combining IMU and other sensing modalities, to obtain multi-modal features that are not highly correlated.

## 6.3 Other Observations

*6.3.1 Blocks and Therapeutic Effects.* From the correlation analysis, we found Diff of both biomarker and behavioral stress are negatively correlated to Before value for children over 20 minutes of play. Does this indicate block play's effect on stress reduction? More long-term follow-up studies and rigorous analysis are needed to answer this question as well as to verify block play's potential therapeutic benefits. For now, we feel our results are sufficient to warrant further research.

Table 8. Activities need to be captured by blocks

| Categories | Play Actions |
|---|---|
| **Individual State** | stand up |
| | lay down |
| **Hand-to-block Interaction** | hold |
| | move |
| | shake |
| **Block-to-block Interaction** | stack |
| | flat |
| | new location |
| **Disassembly** | disassemble |
| | free fall |
| | destroy |

*6.3.2 Behavior Pattern.* From our on-site and video-based observations, some unusual patterns were observed among children who had tsunami exposure and an increased sAA. One such pattern was a behavior of stacking followed by destroying followed by play restricted to a flat surface ("playing flat") (Coastal group P1). This pattern can be related to the “passive play” observed from block features which seemed to indicate a high risk for PTSD. We also discovered that some children who witnessed the tsunami and increased their sAA showed fence building behaviors (Coastal group P14) (Figure 1b). Some aggressive and destructive behaviors such as flicking the blocks was also observed (Figure 1d) among children who did not witness the tsunami (Inland group P5). Among children who increased sAA, many were observed to lack concentration, confirming the negative correlation between the human-annotated play behavior Concentrated Time and sAA Before measurement. Non-concentrated behaviors, including leaving the blocks aside to explore the room, getting attracted by others nearby, as well as tense or passive body movement were observed. In contrast, concentrated children had many trial-and-error sessions and became more engaged with construction if their structures collapsed.

While the above play patterns, such as fence-building, stack-collapse–flat play, stack–collapse-stack play cannot be formally analyzed using our current play features, they could be captured using sequences of block actions, and the relationship between stress and these patterns should be further examined.

## 6.4 Next Steps

*6.4.1 Blocks Design.* Based on our preliminary study results and our observation of specific and unique patterns, we propose a new block design framework to capture several categories of data (Table 8). The first type is individual block states. This type captures the block side that a child puts down, such as stand up (small side on bottom) or lay down (big side on bottom). The second is hand-to-block interactions, including holding, moving, and shaking. The third is the interaction between blocks in a group, in categories such as stacking, play flat, and start a new location. The final category is the manner of disassembly, including normal, free fall, and aggressive destruction. It is necessary to output these data in counts, time, and sequences. In order to accurately detect actions reflecting block surface connections and involving subtle movement, we propose enhanced sensing methods deployed on the surface of the blocks. Utilizing IMU and the screen of a Smartwatch as described in [72], motion data and surface connection can be integrated to detect surface contact-related actions such as Hold, Stack, Flat and Disassemble.

*6.4.2 Experiment Design and New Assessments.* We also propose investigating other mental health targets such as aggressive behavior, depression, and attention, with AssessBlocks. With enhanced block features and more data, we are eager to explore different kinds of predictive methods to improve reliability and the power of assessment.

## 7 CONCLUSION

In the wake of the 2011 Great East Japan Earthquake and Tsunami we designed sensor-embedded toy blocks, to support the automated assessment of young children's mental states after the disaster. We gathered datasets from field studies where children played with our sensor-embedded toy blocks, providing data through a protocol we created. This study provides a positive preliminary assessment of the possibilities of sensor-embedded blocks, for connecting with children exposed to trauma. The study and its results show great potential, but also reveal challenges in both TUI design and data collection. We believe that the potential of our approach, the need for work in this area, and the struggles and limitations we faced when working with those who experienced a natural disaster, can serve future researchers hoping to build a better system to engage health challenges in children. We hope our study promotes child welfare and inspires researches that strive for it.

## ACKNOWLEDGMENTS

This work is supported in part by Japan Society for the Promotion of Science (JSPS) Kakenhi (15K12177, 15K04139) and Natural Sciences and Engineering Research Council of Canada (NSERC) RGPAS/00077-2019 and RGPIN/312218-2013. We would also like to acknowledge Prof. Hiroyasu Kanetaka, Graduate School of Dentistry, Tohoku University, Prof. Yoshie Ohashi, Faculty of Human Welfare, Seigakuin University, Yoshiki Hosoi, Hiromi Sato and Miteki Ishikawa, alumni of Interactive Content Design Lab, Tohoku University, for their contribution to the field study, data collection and analysis.

## REFERENCES

[1] Tomoaki Adachi, Yoshifumi Kitamura, Kazuki Takashima, Toshiki Hosoi, Yoshie Ohashi, Yuichi Ito, and Hiroyasu Kanetaka. 2014. Effect of playing with building blocks on young children with and without posttraumatic stress disorder [in Japanese]. *Miyagi Gakuin Women's University Journal of Developmental Science* 14, 25–30. https://ci.nii.ac.jp/naid/110009807819

[2] Tomoaki Adachi, Yoshifumi Kitamura, Kazuki Takashima, and Miteki Ishikawa. 2017. Effects of block play on salivary alpha-amylase activity in children who lived in affected or less affected area by the tsunami. *EFPA 15th European Congress of Psychology 2017*.

[3] Masahiro Ando, Yuichi Itoh, Toshiki Hosoi, Kazuki Takashima, Kosuke Nakajima, and Yoshifumi Kitamura. 2014. StackBlock: block-shaped interface for flexible stacking. In *Adjunct Proceedings of the ACM Symposium on User Interface Software and Technology*. 41–42. https://doi.org/10.1145/2658779.2659104

[4] ATR-Promotions. 2017. TSND121/151. http://www.atr-p.com/products/TSND121.html

[5] Jennifer Baggerly and Herbert A. Exum. 2007. Counseling children after natural disasters: guidance for family therapists. *The American Journal of Family Therapy* 36, 1, 79–93. https://doi.org/10.1080/01926180601057598

[6] Greg Barish, Eric Elbogen, Patricia Lester, and William R. Saltzman. 2014. Beyond sensors: reading patients through caregivers and context. In *Adjunct Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 1273–1277. https://doi.org/10.1145/2638728.2638811

[7] Susan M Becker. 2007. Psychosocial care for adult and child survivors of the tsunami disaster in India. *Journal of Child and Adolescent Psychiatric Nursing* 20, 3, 148–155. https://doi.org/10.1111/j.1744-6171.2007.00105.x

[8] Judite Blanc, Eric Bui, Yoram Mouchenik, Daniel Derivois, and Philippe Birmes. 2015. Prevalence of post-traumatic stress disorder and depression in two groups of children one year after the January 2010 earthquake in Haiti. *Journal of Affective Disorders* 172, 121–126. https://doi.org/10.1016/j.jad.2014.09.055

[9] Ronald L Blount, Vicki Bunke, Lindsey L Cohen, and Cally J Forbes. 2001. The child–adult medical procedure interaction scale-short form (CAMPIS-SF): validation of a rating scale for children's and adults' behaviors during painful medical procedures. *Journal of Pain and Symptom Management* 22, 1, 591 – 599. https://doi.org/10.1016/S0885-3924(01)00303-7

[10] C. Botella, R. Baños, B. Rey, M. Alcañiz, V. Guillen, S. Quero, and A. García-Palacios. 2006. Using an adaptive display for the treatment of emotional disorders: a preliminary analysis of effectiveness. In *Extended Abstracts on SIGCHI Conference on Human Factors in Computing Systems*. 586–591. https://doi.org/10.1145/1125451.1125574

[11] Barbara Lopes Cardozo, Alfredo Vergara, Ferid Agani, and Carol A. Gotway. 2000. Mental health, social functioning, and attitudes of Kosovar Albanians following the war in Kosovo. *JAMA* 284, 5 (08 2000), 569–577. https://doi.org/10.1001/jama.284.5.569 arXiv:https://jamanetwork.com/journals/jama/articlepdf/192943/joc00483.pdf

[12] Sally Cartwright. 1988. Play can be the building blocks of learning. *Young Children* 43, 5, 44–47. http://www.imaginationplayground.com/images/content/2/9/2987/play-can-be-the-building-blocks-of-learning.pdf

[13] Beth M. Casey, Elizabeth E. Pezaris, and Julie Bassi. 2012. Adolescent boys' and girls' block constructions differ in structural balance: A block-building characteristic related to math achievement. *Learning and Individual Differences* 22, 1, 25–36. https://doi.org/10.1016/J.LINDIF.2011.11.008

[14] Center for research on the epidemiology of disasters. 2018. The human cost of natural disasters: a global perspective. https://www.preventionweb.net/publications/view/42895

[15] Claude M Chemtob, Yoko Nomura, and Robert A Abramovitz. 2008. Impact of conjoined exposure to the World Trade Center Attacks and to other traumatic events on the behavioral problems of preschool children. *Archives of Pediatrics & Adolescent Medicine* 162, 2, 126. https://doi.org/10.1001/archpediatrics.2007.36

[16] Sheldon Cohen, Tom Kamarck, and Robin Mermelstein. 1983. A global measure of perceived stress. *Journal of Health and Social Behavior*, 385–396.

[17] Elysia Poggi Davis and Douglas A Granger. 2009. Developmental differences in infant salivary alpha-amylase and cortisol responses to stress. *Psychoneuroendocrinology* 34, 6, 795–804.

[18] Márcio José Possari dos Santos, Daniel Galera Bernabé, Ana Cláudia de Melo Stevanato Nakamune, Silvia Helena Venturoli Perri, Sandra Helena Penha de Oliveira, et al. 2012. Salivary alpha amylase and cortisol levels in children with global developmental delay and their relation with the expectation of dental care and behavior during the intervention. *Research in developmental disabilities* 33, 2, 499–505.

[19] Charles H Elliott, Susan M Jay, and Patricia Woody. 1987. An observation scale for measuring children's distress during medical procedures. *Journal of Pediatric Psychology* 12, 4, 543–551. https://doi.org/10.1093/jpepsy/12.4.543

[20] Min Fan, Alissa N Antle, Maureen Hoskyn, Carman Neustaedter, and Emily S Cramer. 2017. Why tangibility matters: a design case study of at-risk children learning to read and spell. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1805–1816. https://doi.org/10.1145/3025453.3026048

[21] Ruth Feldman, Adva Vengrober, Moranne Eidelman-Rothman, and Orna Zagoory-Sharon. 2013. Stress reactivity in war-exposed young children with and without posttraumatic stress disorder: Relations to maternal stress hormones, parenting, and child emotionality and regulation. *Development and Psychopathology* 25, 4pt1, 943–955. https://doi.org/10.1017/S0954579413000291

[22] Pietro Ferrara, Giorgia Bottaro, Silvia Angeletti, Antonio Gatto, Ottavio Vitelli, Domenica Battaglia, Maria Del Re, Antonio Ruggiero, and Giordano Dicuonzo. 2013. Salivary alpha-amylase: a new non-invasive biomarker for assessment of pain perception in epileptic children. *Acta Neurologica Belgica* 113, 3, 279–283.

[23] Davide Figo, Pedro C. Diniz, Diogo R. Ferreira, and João M. P. Cardoso. 2010. Preprocessing techniques for context recognition from accelerometer data. *Personal and Ubiquitous Computing* 14, 7, 645–662. https://doi.org/10.1007/s00779-010-0293-9

[24] NF Furlan, MBD Gaviäo, TS Barbosa, J Nicolau, and Paula Midori Castelo. 2012. Salivary cortisol, alpha-amylase and heart rate variation in response to dental treatment in children. *Journal of Clinical Pediatric Dentistry* 37, 1, 83–87.

[25] Audrey Girouard, David McGookin, Peter Bennett, Orit Shaer, Katie A. Siek, and Marilyn Lennon. 2016. Tangibles for health workshop. In *Extended Abstracts of the SIGCHI Conference on Human Factors in Computing Systems*. 3461–3468. https://doi.org/10.1145/2851581.2856469

[26] Roberta A. Goodfader. 1982. Sex differences in the play constructions of pre-school children. *Smith College Studies in Social Work* 52, 2, 129–144. https://doi.org/10.1080/00377318209516556

[27] Elisabeth S. Hirsch. 1996. *The block book*. National Association for the Education of Young Children.

[28] Thomas L. Holzer and James C. Savage. 2013. Global earthquake fatalities and population. *Earthquake Spectra* 29, 1, 155–175. https://doi.org/10.1193/1.4000106

[29] Anthony Hornof, Haley Whitman, Marah Sutherland, Samuel Gerendasy, and Joanna McGrenere. 2017. Designing for the "Universe of One": personalized interactive media systems for people with the severe cognitive impairment associated with Rett syndrome. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2137–2148. https://doi.org/10.1145/3025453.3025904

[30] Toshiki Hosoi, Kazuki Takashima, Tomoaki Adachi, Yuichi Itoh, and Yoshifumi Kitamura. 2014. A-blocks: recognizing and assessing child building processes during play with toy blocks. In *SIGGRAPH Asia 2014 Emerging Technologies*. 1–2. https://doi.org/10.1145/2669047.2669061

[31] Jittrapol Intarasirisawat, Chee Siang Ang, Christos Efstratiou, Luke William Feidhlim Dickens, and Rupert Page. 2019. Exploring the touch and motion features in game-based cognitive assessments. In *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol. 3. Article 87, 25 pages. https://doi.org/10.1145/3351245

[32] Akemi Ishigaki, Hikari Higashi, Takako Sakamoto, and Shigeki Shibahara. 2013. The great east-Japan earthquake and devastating tsunami: an update and lessons from the past great earthquakes in Japan since 1923. *The Tohoku Journal of Experimental Medicine* 229, 4, 287–299. https://doi.org/10.1620/tjem.229.287

[33] Sigal Jacoby, Galia Gutwillig, Doron Jacoby, Naomi Josman, Patrice L. Weiss, Minori Koike, Yuichi Itoh, Norifumi Kawai, Yoshifumi Kitamura, and Ehud Sharlin. 2009. PlayCubes: monitoring constructional ability in children using a tangible user interface and a playful virtual environment. In *Proceedings of the IEEE Virtual Rehabilitation International Conference*. 42–49. https://doi.org/10.1109/ICVR.2009.5174203

[34] Japan Meteorological Agency. 2020. Number of aftershocks of M5.0 or more [in Japanese]. http://www.data.jma.go.jp/svd/eqev/data/2011{_}03{_}11{_}tohoku/aftershock/

[35] Zhaobao Jia, Wenhua Tian, Xiang He, Weizhi Liu, Chunlin Jin, and Hansheng Ding. 2010. Mental health and quality of life survey among child survivors of the 2008 Sichuan earthquake. *Quality of Life Research* 19, 9, 1381–1391. https://doi.org/10.1007/s11136-010-9703-8

[36] Heidi Kaduson and Charles E. Schaefer. 2010. *101 favorite play therapy techniques. Volume III.* Jason Aronson. 430 pages.

[37] Hiroko Kamide, Kazuki Takashima, Miteki Ishikawa, Tomoaki Adachi, and Yoshifumi Kitamura. 2018. Quantitative evaluation of childrenapos;s developmental stage in building block play [in Japanese]. *The Transactions of Human Interface Society* 20, 1, 107–114. https://doi.org/10.11184/his.20.1_107

[38] Abiola Keller, Kristin Litzelman, Lauren E. Wisk, Torsheika Maddox, Erika Rose Cheng, Paul D. Creswell, and Whitney P. Witt. 2012. Does the perception that stress affects health matter? the association with health and mortality. *Health Psychology* 31, 5, 677–684. https://doi.org/10.1037/a0026743

[39] A M Khan, Young-Koo Lee, S Y Lee, and Tae-Seong Kim. 2010. A triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical recognizer. *IEEE Transactions on Information Technology in Biomedicine* 14, 5, 1166–1172. https://doi.org/10.1109/TITB.2010.2051955

[40] Naoko Kimura. 2009. How to do the screening for developmental disorder in the routine 18 month and 36 month health check up [in Japanese]. *Research bulletin of Naruto University of Education* 24, 13–19. https://ci.nii.ac.jp/els/contentscinii{_}20180108082303.pdf?id=ART0008984834

[41] Yoshifumi Kitamura, Yuichi Itoh, and Fumio Kishino. 2001. Real-time 3D interaction with ActiveCube. In *Extended Abstracts on the SIGCHI Conference on Human Factors in Computing Systems*. 355. https://doi.org/10.1145/634067.634277

[42] Kalsea J Koss, Melissa RW George, E Mark Cummings, Patrick T Davies, Mona El-Sheikh, and Dante Cicchetti. 2014. Asymmetry in children's salivary cortisol and alpha-amylase in the context of marital conflict: Links to children's emotional security and adjustment. *Developmental Psychobiology* 56, 4, 836–849.

[43] Carolyn Kousky. 2016. Impacts of natural disasters on children. *The Future of Children* 26, 1, 73–92. http://www.jstor.org/stable/43755231

[44] Brian M. Landry, Eun Kyoung Choe, Stephen McCutcheon, and Julie A. Kientz. 2010. Post-traumatic stress disorder: opportunities challenges for computing technology. In *Proceedings of the ACM International Health Informatics Symposium*. 780–789. https://doi.org/10.1145/1882992.1883110

[45] Ye Liu, Xinye Zhang, Jinshi Cui, Chen Wu, Hamid Aghajan, and Hongbin Zha. 2010. Visual analysis of child-adult interactive behaviors in video sequences. In *Proceedings of the IEEE International Conference on Virtual Systems and Multimedia*. 26–33. https://doi.org/10.1109/VSMM.2010.5665969

[46] John Mirowsky. 1999. *Analyzing associations between mental health and social circumstances.* 105–123. https://doi.org/10.1007/0-387-36223-1_6

[47] Elizabeth Mullett-Hume, Daphne Anshel, Vivianne Guevara, and Marylene Cloitre. 2008. Cumulative trauma and posttraumatic stress disorder among children exposed to the 9/11 World Trade Center attack. *American Journal of Orthopsychiatry* 78, 1, 103–108. https://doi.org/10.1037/0002-9432.78.1.103

[48] Urs M. Nater, Nicolas Rohleder, Jens Gaab, Simona Berger, Andreas Jud, Clemens Kirschbaum, and Ulrike Ehlert. 2005. Human salivary alpha-amylase reactivity in a psychosocial stress paradigm. *International Journal of Psychophysiology* 55, 3, 333–342. https://doi.org/10.1016/j.ijpsycho.2004.09.009 arXiv:NIHMS150003

[49] Nichigan Original. 2017. Unpainted Wooden Blocks. http://www.nocorp.co.jp/mutoso/

[50] Mary Jo. Pollman. 2010. *Blocks and beyond : strengthening early math and science skills through spatial learning.* Paul H. Brookes Pub. Co. 176 pages.

[51] Tara M. Powell and Tuyen Bui. 2016. Supporting social and emotional skills after a disaster: findings from a mixed methods study. *School Mental Health* 8, 1, 106–119. https://doi.org/10.1007/s12310-016-9180-5

[52] R S Pynoos, A Goenjian, M Tashjian, M Karakashian, R Manjikian, G Manoukian, A M Steinberg, and L A Fairbanks. 1993. Post-traumatic stress reactions in children after the 1988 Armenian earthquake. *The British Journal of Psychiatry : The Journal of Mental Science* 163, 2, 239–47. https://doi.org/10.1192/BJP.163.2.239

[53] Linda A. Reddy, Tara M. Files-Hall, and Charles E. Schaefer. 2005. *Empirically Based Play Interventions for Children.* American Psychological Association, Washington. https://doi.org/10.1037/11086-000

[54] Albert Rizzo, Arno Hartholt, Mario Grimani, Andrew Leeds, and Matt Liewer. 2014. Virtual reality exposure therapy for combat-related posttraumatic stress disorder. *Computer* 47, 7, 31–37. https://doi.org/10.1109/MC.2014.199

[55] T. Sadoh and K. Nakato. 1987. Surface properties of wood in physical and sensory aspects. *Wood Science and Technology* 21, 2, 111–120. https://doi.org/10.1007/BF00376191

[56] Gopal K Sahu, Seema Upadhyay, and Shradha M Panna. 2014. Salivary alpha amylase activity in human beings of different age groups subjected to psychological stress. *Indian journal of clinical biochemistry : IJCB* 29, 4, 485–490. https://doi.org/10.1007/s12291-013-0388-y
[57] Benedetto Saraceno. 2005. The WHO's mental health response to the Asian tsunami. *World Psychiatry : Official Journal of the World Psychiatric Association* 4, 2, 66–7.
[58] Julie Sarama and Douglas H. Clements. 2009. Building blocks and cognitive building blocks: playing to know the world mathematically. In *American Journal of Play*, Vol. 1. 313–337. https://eric.ed.gov/?id=EJ1069014
[59] Farhana Shahid, Wasifur Rahman, Anika Binte Islam, Nipi Paul, Nabila Khan, Mohammad Saifur Rahman, Md Munirul Haque, and A. B. M. Alim Al Islam. 2019. Two tell-tale perspectives of PTSD: neurobiological abnormalities and Bayesian regulatory network of the underlying disorder in a refugee Context. In *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol. 3. ACM, New York, NY, USA, Article 101, 45 pages. https://doi.org/10.1145/3351259
[60] Ehud Sharlin, Yuichi Itoh, Benjamin Watson, Yoshifumi Kitamura, Steve Sutphen, and Lili Liu. 2002. Cognitive Cubes: A tangible user interface for cognitive assessment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 347–354. https://doi.org/10.1145/503376.503438
[61] Yukiko Shimomura, Mikae Fukasawa, and Kazunori Takeda. 2010. Stress evaluation in children with cancer undergoing invasive medical procedures [in Japanese]. *Journal of Life Support Engineering* 22, 3, 112–118. https://doi.org/10.5136/lifesupport.22.112
[62] Beverly A. Snyder. 1997. Expressive art therapy techniques: healing the soul through creativity. *The Journal of Humanistic Education and Development* 36, 2, 74–82. https://doi.org/10.1002/j.2164-4683.1997.tb00375.x
[63] Katharina Spiel, Christopher Frauenberger, Eva Hornecker, and Geraldine Fitzpatrick. 2017. When empathy is not enough: assessing the experiences of autistic children with technologies. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2853–2864. https://doi.org/10.1145/3025453.3025785
[64] Zachary Steel, Derrick M Silove, Tuong Phan, and Adrian Bauman. 2002. Long-term effect of psychological trauma on the mental health of Vietnamese refugees resettled in Australia: a population-based study. *The Lancet* 360, 1056–1062.
[65] Takashi Sugimoto, Tomohiro Shinozaki, and Yuki Miyamoto. 2013. Aftershocks associated with impaired health caused by the great east Japan disaster among youth across Japan: A national cross-sectional survey. *Journal of Medical Internet Research* 15, 12. https://doi.org/10.2196/ijmr.2585
[66] S. A. Taylor, N. Jaques, E. Nosakhare, A. Sano, and R. Picard. 2017. Personalized multitask learning for predicting tomorrow's mood, stress, and health. *IEEE Transactions on Affective Computing*, 1–1. https://doi.org/10.1109/TAFFC.2017.2784832
[67] Hideki Tsumura, Hironori Shimada, Hiroshi Morimoto, Chihiro Hinuma, and Yoshiko Kawano. 2014. Effects of distraction on negative behaviors and salivary $\alpha$-amylase under mildly stressful medical procedures for brief inpatient children. *Journal of health psychology* 19, 8, 1079–1088.
[68] UNISDR. 2018. Disaster Statistics. https://www.unisdr.org/we/inform/disaster-statistics
[69] VAS. 2017. Visual Analogue Scale - Physiopedia. https://www.physio-pedia.com/Visual{_}Analogue{_}Scale
[70] Emanuel Vonach, Marianne Ternek, Georg Gerstweiler, and Hannes Kaufmann. 2016. Design of a health monitoring toy for children. *Proceedings of the International Conference on Interaction Design and Children*, 58–67. https://doi.org/10.1145/2930674.2930694
[71] Ping Wang, Gregory D. Abowd, and James M. Rehg. 2009. Quasi-periodic event analysis for social game retrieval. In *Proceedings of IEEE International Conference on Computer Vision*. 112–119. https://doi.org/10.1109/ICCV.2009.5459151
[72] Xiyue Wang, Miteki Ishikawa, Kazuki Takashima, Tomoaki Adachi, Ehud Sharlin, Patrick Finn, and Yoshifumi Kitamura. 2018. Designing action-characterizing toy blocks for behavior assessments. In *Extended Abstracts of the SIGCHI Conference on Human Factors in Computing Systems*. Article Paper LBW513, 6 pages. https://doi.org/10.1145/3170427.3188451
[73] Tracy L. Westeyn, Gregory D. Abowd, Thad E. Starner, Jeremy M. Johnson, Peter W. Presti, and Kimberly A. Weaver. 2012. Monitoring children's developmental progress using augmented toys and activity recognition. *Personal and Ubiquitous Computing* 16, 2, 169–191. https://doi.org/10.1007/s00779-011-0386-0
[74] WHO. 2012. Mental Health Assistance to the Populations Affected by the Tsunami in Asia. http://www.who.int/mental{_}health/resources/tsunami/en/index1.htmlpapers2://publication/uuid/4836616E-37B7-45C6-887D-7F226322CC9D
[75] Naomi Yamashita, Hideaki Kuzuoka, Keiji Hirata, Takashi Kudo, Eiji Aramaki, and Kazuki Hattori. 2017. Changing moods: how manual tracking by family caregivers improves caring and family communication. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 158–169. https://doi.org/10.1145/3025453.3025843
[76] Bingshu Yang, Jinshi Cui, Hongbin Zha, and Hamid Aghajan. 2012. Visual context based infant activity analysis. In *IEEE International Conference on Distributed Smart Cameras*. 1–6. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6470169&isnumber=6470120