Analysis of cognitive states during bodily exploration of mathematical concepts in visually impaired children

Erica Volta
Casa Paganini - InfoMus, DIBRIS
University of Genoa
Genoa, Italy
erica.volta@edu.unige.it

Radoslaw Niewiadomski
CONTACT Unit
Istituto Italiano di Tecnologia
Genoa, Italy
radoslaw.niewiadomski@iit.it

Temitayo Olugbade
UCLIC
University College of London
London, UK
temitayo.olugbade.13@ucl.ac.uk

Carla Gilio
Istituto Chiossone
Genoa, Italy
gilio@chiossone.it

Elena Cocchi
Istituto Chiossone
Genoa, Italy
cocchi@chiossone.it

Nadia Berthouze
UCLIC
University College of London
London, UK
n.berthouze@ucl.ac.uk

Monica Gori
U-VIP
Istituto Italiano di Tecnologia
Genoa, Italy
monica.gori@iit.it

Gualtiero Volpe
Casa Paganini - InfoMus, DIBRIS
University of Genoa
Genoa, Italy
gualtiero.volpe@unige.it

***Abstract*—When developing interactive systems for children, such as serious games in the context of educational technology, it is important to take into account and address relevant cognitive and emotional child's experiences that may influence learning outcomes. Some works were done to analyze and automatically recognize these cognitive and affective states from nonverbal expressive behaviors. However, there is a lack of knowledge about visually impaired children and their body language to convey those states during learning tasks. In this paper, we present an analysis of nonverbal expressive behaviors of both blind and low-vision children, aiming at understanding what type of body communication can be an indicator of two cognitive states: engagement and confidence. In the study we consider the data collected along the EU-ICT H2020 weDRAW Project, while children were asked to solve mathematical tasks with their body. For such a dataset, we propose a list of 31 nonverbal behaviors, annotated both by rehabilitators used to work with visually impaired children and by naive observers. In the last part of the paper, we propose a preliminary study on automatic recognition of engagement and confidence states from 2D positional data. The classification results are up to 0.71 (F-score) on a three-class classification task.**

***Index Terms*—visually impaired children, engagement, confidence, learning, cognition, nonverbal behaviors, classification, machine learning**

I. Introduction

Nonverbal communication is the first stage of communication development [41], representing a fundamental part of what people use to convey information when they interact and communicate with each other [23]. Since the visual modality is crucial for developing social and communication abilities, visually impaired children may show nonverbal behaviors different from children without visual impairment [46]. In this paper, we focus on nonverbal indicators of cognitive states in blind children and children with visual impairment in mathematical task-oriented settings. The work is carried-out within the EU-H2020-ICT weDRAW Project that aimed at developing adaptive multisensory technologies to enhance understanding of mathematical concepts in primary school children. According to the project objectives, a personalized technology design can meet the needs of both typically developed and visual- or learning- impaired (e.g., blind, or dyslexic) children. In this framework, a set of serious games has been developed to teach children the basic concepts of mathematics. Such games would benefit from an adaptive feedback based on a child's specific learning needs and her/his involvement in the task, e.g., by means of an automatic measurement of her/his cognitive engagement and confidence. Whilst several works exist on the recognition of cognitive states (e.g., see [27], [36] for engagement) they do not focus on nonverbal behaviors of visually impaired users. In Sections II, III, IV, and V, starting from the state-of-the-art on nonverbal communication in visually impaired children and from the weDRAW dataset, we describe relevant nonverbal cues of engagement and confidence for this population. We then introduce the annotation procedure we adopted in this study in Section VI, and we present our preliminary work on automatic classification of these states from video in Section VII.

The weDRAW Project has received funding from the European Union's Horizon 2020 Research and Innovation Program under Grant Agreement No.732391.

II. Background: gestures and unconstrained body movement in visually impaired children

The level of visual functioning can greatly influence early child development. A visual disability may therefore lead to developmental delays, especially if an early intervention does not take place [49]. Developmental delays can have a negative effect on child's participation both in rehabilitation and educa-

978-1-7281-3891-6/19/$31.00 ©2019 IEEE

Fig. 1: Picture of a blind child exploring angles with his arms during the weDRAW data collection

tional settings. In addition, poorer immediate problem-solving [13] and mathematical skills [30] have been found in children with a visual impairment.

Nonverbal behaviors in visually impaired children were expected to be semantically different from those of typically developing ones. For this reason, considering the state-of-the-art, it is relevant to investigate the extent to which visual impairment may affect the development of nonverbal communication patterns and the ability to produce such patterns in various social interactions. For instance, head orientation, proxemics, and distance from objects (e.g., walls) might have different meaning [46] when displayed by visually impaired individuals. Iverson and Golden-Meadow [26], for example, discussed gestures used by congenitally blind children who never saw gestures before nor experienced their communicative functions. Results showed that visually impaired children produced gestures, but not in all situations as it was for sighted and blindfolded ones. The study suggested that gestures provide speaker with functions that are independent by the listeners.

III. Engagement and confidence behaviors in children

Important developmental and social changes occur in children starting from the age of 6. Through these years, they contribute to create a personal identity, a self-concept, and an orientation toward achievements that will be relevant for shaping their success in school, work, and life. In this work, we investigate confidence and engagement in primary school children with visual impairments during embodied mathematical tasks solving. The educational literature shows (see [38] for a review) that in interactive learning task-oriented environments, it is crucial to be able to recognize these two cognitive states in order to avoid the potentially negative outcomes of learning experience. Lack of engagement, or low self-confidence might lead to negative emotions such as boredom or anxiety [9] and even to the complete abandonment of the task. Below, we provide the definitions of the two states as they were given to the annotators:

Engagement can be considered a cognitive construct, based on the interrelation between behavior, cognition and emotion. In young children, it can be deduced from a child's interaction, other peers and materials in a way that is appropriate from a developmental and contextual point of view [31], [48]. In learning studies, engagement is often defined as the amount of energy that the student devotes to learning experience. In literature, research highlighted three different types of students engagement: cognitive, emotional, and behavioral [18]. Cognitive engagement is related to how much students invest cognitive efforts and resources in learning; the emotional one consider motivation and commitment; instead behavioral engagement deals with students on-task behaviors. [35].

On the other hand, *confidence* results from the appreciation of one's own abilities or qualities. Scientific literature [14], [40] has corroborated Erikson's idea [15] that feelings of competence and personal esteem are of central importance to a child's well-being [25]. For example, children who do not consider themselves competent in academic, social, or other fields (such as athletics, music, theatre or scouting) during their elementary school years report depression and social isolation more often than their peers [12]. Self-confidence is closely related to the task for which the solution is being sought and can also be observed in short intervals of time.

Cognitive engagement has a great impact on learning outcomes and this is why it is such considered also in e-learning technology research [16]. Analogously, self-confidence is one of the states that teachers use to monitor children learning outcomes [20].

IV. Automatic recognition of engagement and confidence

Several works were proposed in the literature to computationally address the level of engagement from nonverbal behaviors [29]. Most of them focused on engagement detection or estimation in human-human [17], human-virtual agent [33], or human-robot [3] dyadic or group interactions from gaze [33], back-channels [43], and facial expressions [21]. Frank and colleagues [17] proposed the Engagement Classification Framework composed of six states from "disengagement" to "involved action". Their framework implementation detects three levels of engagement from 3D data by detecting a set of features of the upper body movements such as hand vertical positions and speed, leaning and body direction as well as specific postures. In [28], a multimodal approach was proposed to detect levels of engagement, using nonverbal features extracted from audio and visual data, and using rank learning. Finally, recent works included the application of deep learning techniques to compute intensity of engagement from video in e-learning tasks (see, e.g., papers included in the "Engagement in the Wild track" of the EmotiW 2018 Challenge at the ICMI conference[1]).

In the context of single-user activity, Ge and colleagues [19] proposed a model for engagement/disengagement detection in

[1] https://sites.google.com/view/emotiw2018

autistic children from body movements. They detected each child's concentration on a given task (e.g., playing a game on a tablet) by extracting a set of features from kinematics data obtained from an RGB-D camera and applying machine learning techniques. Several features were computed using angles and distances between child's joints and an object of interest (e.g., a table). Using standard algorithms such as SVM, Random Forest, and AdaBoost, they achieved a recognition rate up to 97% for a two-class pattern recognition problem. Shaker and Shaker [45] detected the level of engagement from nonverbal cues in a context of single-user video game. They extracted several low-level visual features, and combined them with high-level facial expressions labels. Next, they applied Neuroevolutionary Preference Learning (NPL) to obtain an accuracy of 96%.

Computational approaches to confidence level are more rare. Most frequently, researchers focused on similar topics such as leadership detection in multi-user scenarios, e.g., social games [4], and self-efficacy in s pain-related scenario [37]. We are unaware, at this stage, of any existing model for the recognition of confidence level from full body cues in the context of a single-user task in education.

V. The VI-weDRAW Dataset

Body movement plays multiple roles in the weDRAW project: first, it is a means both for the child and the teacher to explore, construct, and understand some arithmetical and geometrical concepts [38]; second: it allows an observer (either the teacher or the technology) to gain insights on a child's cognitive and affective processes that affect learning [26]. In particular, in this paper, we only focus on the second aspect, i.e., a communication channel allowing humans to express and perceive implicit high-level messages, such as emotional and cognitive states or social boundaries. Thus, it is out of the scope of this paper to recognize whether the child was able to perform correctly the task through the appropriate choice of gestures and poses.

A. Dataset and Participant Profile

For the purpose of the work we created, with the participation of visually impaired children, the VI-weDraw dataset, which comprises body movement data, captured during mathematical problem-solving specifically designed for the experimental setting. The tasks were based on project premises and included: angles, symmetry and reflection, considered as a type of isometric transformation of shapes. The dataset consists of two synchronized video recordings (frontal and lateral) with corresponding audio data. All the children recruited for the experiment were studying in Genoa and participating in a rehabilitation programme at Chiossone Institute in Genoa at the time of the data collection. We collected the data from 3 blind and 14 low-vision children. The visual acuity of the collective group ranged from no perception of the light to visual acuity of 1/20 from an eye. To understand the level of cognitive impairment, the verbal QI (QIV) and performance QI (QIP) items of the Weschler Intelligence Scale for Children (WISC-IV) [47] were used as well as the IAG and the Griffiths-III [22] tests. Collectively, the cognitive tests showed that 3 of the 17 children had levels of cognitive impairments.

B. The mathematical problem-solving tasks

Following the procedure described in [39] with typical developing children, the data collection session started with an exploration of static representation of angles using child's arms, whose movements were supported with sonification technology. Sounds were realized using a tonal scale played by strings instruments, and the pitch was mapped to the inner angle (along the vertical plane) between childś arms. It was played in real-time, according to child's arms aperture (see Fig. 1).

Differently from [39], the children were also trained to use proprioceptive and tactile feedback with a flat wall as reference: standing with the back flat against the wall represents 90 degrees angle. The child's arms extended outwards ipsilaterally (and against the wall) to form a 180 degrees angle. By extending one arm outward ipsilaterally and the other contralaterally, the children formed a 0 degrees angle, while extending one arm outward ipsilaterally and the other in the anterior they created a 90 degrees angle. For each child the task was first explained by the instructor, who helped the child's exploration of angles 0 degrees, 90 degrees and 180 degrees as described above. Next, the child was encouraged to explore sounds feedback of her/his own. Finally, the child was asked to represent 45 degrees and 135 degrees angles. In the second task, each child was asked to represent 45 degrees angle (and 135 degrees angle) by rotating his/her body about the vertical axis. This additional proprioceptive and tactile exploration was suggested by the expert in visual disability rehabilitation working in the project.

VI. Annotation and labeling

From the video data, we extracted the episodes in which each child attempt to solve the given mathematical problems. We collected in this way 92 episodes with total duration of 12 minutes and 10 seconds. All the episodes were presented to the annotators in random order and without audio information. Although in the literature there are findings suggesting that for learning related tasks, untrained observers can also provide reliable ratings, we decided firstly to ask to four experts (co-authors of this paper) to annotate each episode. The group was composed of one rehabilitation specialist from Chiossone Institute for visually impaired people, and three experts in movement analysis from University of Genoa. They used a 1-to-5 Likert scale, with 3 as neutral state, for each of the degrees of confidence and engagement Each vote was a result of the informal consensus between all the experts. Meanwhile, they also discussed and jointly reported the nonverbal behaviors observed. The same experts proposed a list of 31 observed nonverbal behaviors (see Tab. I). For the second round of annotations, we chose 20 episodes, from the initial set of 92, which in the above mentioned ranking were annotated as representing high/low engagement and high/low confidence.

TABLE I: The list of nonverbal behaviors and their frequencies of appearance in the dataset

Movement quality			Posture			Gesture		
Id	*Cues*	*%*	*Id*	*Cues*	*%*	*Id*	*Cues*	*%*
1	Focused, direct movement	34.44	10	Gaze down	38.99	21	Exp. of positive emotions (e.g. laughter)	18.33
2	Jerky movement	25.00	11	Tendency to act	25.00	22	Nervous smile or laughter	13.99
3	Hesitating movement	22.78	12	Listening predisposition	25.00	12	Open mouth	10.00
4	Fluid movement	21.11	13	Body as a reference point	20.00	21	Nodding during tasks resolution	7.88
5	Impulsive movement	20.00	14	Gaze contact with the interlocutor	18.33	25	Grabbing clothes	7.88
6	Inhibited movement	17.22	15	Withdraw from action	16.77	26	Rocking	7.88
7	Not goal-oriented movement	16.77	16	Outward-facing gaze	13.99	27	Lips biting	6.77
8	Slow movement	15.00	17	A loss of posture alignment	13.33	28	Deictic gestures	5.00
9	Misalignment of different body planes	11.11	18	A loss of balance (feet support instability)	13.32	29	Hands kept together	5.00
						30	Hands hold behind back	4.44
			19	Posture openness	8.89	31	Touching face or mouth	0
			20	Legs are moved while body is still	7.78			

TABLE II: Cues associated with high-level of engagement annotation, expressed in frequency above median value.

id	Nonverbal Cues	% Annotation Frequency
1	Focused, direct movement	25.56
11	Tendency to act	18.99
12	Listening predisposition	17.22
4	Fluid movement	16.11
21	Expressing positive emotions (e.g. laughter)	15.00

TABLE III: Cues associated with high-level of confidence annotation, expressed in frequency above median value.

id	Nonverbal Cues	% Annotation Frequency
1	Focused, direct movement	31.66
11	Tendency to act	21.77
10	Gaze down	21.11
12	Listening predisposition	15.66
5	Impulsive movement	15.00

TABLE IV: Cues associated with low-level of engagement annotation, expressed in frequency below median value.

id	Nonverbal Cues	% Annotation Frequency
10	Gaze down	19.44
2	Jerky movement	13.89
15	Withdraw from the action	12.78
3	Hesitating movement	10.56
7	Not goal-oriented movement	10.56

TABLE V: Cues associated with low-level of confidence annotation, expressed in frequency below median value.

id	Nonverbal Cues	% Annotation Frequency
10	Gaze down	19.44
3	Hesitating movement	14.44
2	Jerky movement	13.33
6	Inhibited movement	10.00
7	Not goal-oriented movement	10.00

Three experts (one psychomotrician and two specialists in orientation and mobility for visually impaired patients) and six non-experts annotated all 20 episodes displayed in random order. We decided to perform a second annotation by both experts and non experts to check whether particular patterns are differently recognized by these two groups of annotators, meanwhile we expected the same types of observations by the three experts compared to those of the first round annotation. Annotators were asked to indicate relevant nonverbal behaviors from the list presented in Tab. I, and the perceived level of child engagement and confidence (defined as above), using a 1-to-5 Likert scale.

To measure inter-rater agreement on perceived engagement and confidence, we computed two-ways random ICC average agreement [24] of the Likert values. Results are 0.5291 for engagement and 0.4662 for confidence. These ICCs are fair [11], especially considering the mixture of expertise of the annotators.

For annotation of nonverbal behaviors, we firstly computed the frequency with which each cue was selected as relevant. The nonverbal behaviors: 1) related to gaze direction, 2) movements qualities, such as directness, hesitation, or jerkiness, 3) involvement or retraction from the action, and preparation to listen were frequently chosen as relevant (see Tab. I).

To understand the relation between behaviors and the targeted cognitive states, we computed annotation frequency of each behavior in only videos ranked as “high engagement” and “high confidence”. We used the median value for each cognitive state (for engagement equal to 4 and for confidence to 3) as a threshold for separating the episodes rated as “high” from the“low” ones. Next, for each behavior, we computed annotation frequency separately for videos above and below these thresholds. The most frequent behavior for each state is reported in Tab. II (high engagement), and Tab. III (high confidence). As it can be seen, two of them are presented in both states: *focused movement* and *tendency to act*. Thus, they can be indicators of both high engagement and confidence.

We used the same approach to find nonverbal cues of low engagement and confidence. Results in Tab. IV and V show that low engagement and confidence are mostly expressed with *gaze down*, *jerky* and *hesitating movement*, and *non goal-oriented movement*. Interestingly, cues appear frequently for low levels of both cognitive states (e.g., *gaze down*, or *jerky movement*).

As reported in Tab. III- V, some cues were frequently annotated: 1) for both cognitive states, and 2) in high and low level of the same cognitive state. This might give to the reader the impression that these cues are not discriminative. There are, however, possible explanations for this observation. Regarding case 1), it is possible that such cues, e.g. *focused movement*, are relevant for a cognitive state only if they co-occur with some other behaviors. The sharing of specific cues between low confidence and engagement may also be explained by the fact that the two states often co-occur. Indeed frequently, low level of self-confidence tends to influence engagement and participation in task, especially considering the learning tasks and children's young ages [14]. Regarding case 2, it is important to notice that some of the cues are binary (e.g., *appearance of laughter as expression of positive emotion*) whilst others can be considered as continuous cues (e.g., *fluid movement)*). In the case of continuous cues, it might happen that different "degrees" of such a cue can be associated with different levels of the corresponding cognitive state. This example shows a shortcoming of our annotation schema, as so far we have not used a continuous scale for the annotation of nonverbal behaviors. Thus, future works are needed to address this shortcoming.

Considering these preliminary results and the final aim of the work, it is interesting to highlight that cues as gaze, generally considered as one of the fundamental cues in video detection of children and students engagement [42], [44], had a relevant role for human annotators also in the context of visually impaired people, especially in the case of low engagement and confidence. Literature on early-social cognitive development, deeply analyzed the use of gaze, starting from infancy, as a privileged cue of social interaction and others' attention and intention. When *mutual gaze* occurs, according to theorists [32], it is a sign of social engagement and mutual interaction, whilst *following gaze* is considered a sign of understanding others' attention. *Gaze alternation*, in dyadic or group interaction, is used to assess joint attention [1]. Those cues were largely considered, for example, in understanding engagement in autistic children [10]. From the literature, we know that blind children have difficulties in detecting patterns of social interactions, meanwhile sighted people surrounding them may have difficulties assessing where a blind child focuses her attention, since there is neither visual orienting and pointing, and gazing and facial expressions are more neutral [5]. For this reason, we may suppose that the absence of such common patterns of joint attention and engagement, led annotators (who were in majority non-experts), to consider still position of gaze, looking down, as a cue of both lack of engagement and self-confidence, as this is how it is perceived by typically developed people in social interaction contexts. To check this hypothesis (and be able to perform a comparative analysis between non-experts and experts), we need, however, to collect more annotation from experts.

VII. Automated classification

A. Features extraction

To check whether it is possible to detect the levels of engagement and confidence from visual data in the context of a single-user task, we performed a series of preliminary experiments with standard machine learning techniques on the dataset presented in Section V. Due to the small number of annotated episodes, we subdivided them into smaller segments of fixed duration of 50 frames (corresponding to 1 second), obtaining 758 segments. The labels for each of the two states were not balanced. The same shortcoming was observed after the sub-segmentation. Thus, we decided to regroup the rates:

- Levels 4 and 5 on the engagement scale were regrouped into high level engagement,
- Levels 1 and 2 on the confidence scale were regrouped into low engagement level.
- Level 3 expressed medium level of engagement.

We obtained in this way, 294 segments for high engagement, 176 for medium and 288 for low engagement and 172 segments for high confidence, 317 for medium and 269 segments for low confidence. 15 features were extracted from 2D positional data obtained by applying OpenPose [7] to the frontal view recordings. The features were computed on: 1) front head, 2) left and right elbow, 3) left and right knee, 4) Cervical vertebrae (C7). Eight features are low-level kinematic features: *Right* and *Left Arm Position Variances*, *Velocities*, *Kinetic Energies*, *Head Velocity* and *Kinetic Energy*. *Head Side Leaning* is computed as a difference between the x coordinate of the head and the mean of the x coordinates of the upper limbs. The algorithm by [34] is used to compute *Right Arm and Left Arm Stability*, [2] to compute *Arms Fluidity Index* and [8] to compute *Closure Area*. *Body Symmetry Index* is the sum of: 1) the sum of absolute differences between x (resp. y) coordinates of the upper body limbs and C7, 2) the absolute difference of the head and C7 x coordinates. Finally, *Foot Symmetry* is computed as the distance between the x coordinates of the child lower body limbs and C7 x coordinate. It is important to state that, as a first step, these features cover only a subset of the cues listed in Tab. III-V. The kinematic features are low-level components of various expressive qualities (first column of Tab. I), such as fluidity or impulsivity [6]. The remaining features correspond to some of the observed postures and gestures (second and third column of Tab. I), e.g., loss of balance (19), posture openness (20).

Next, three aggregation operators (average, maximum, and minimum) were applied to the values computed on the single observations. Thus, a 45-feature vector was used for each segment. Finally, the data was normalized using the z-normalization method.

B. Classification

We performed a set of machine learning experiments to automatically classify three levels of confidence and engagement (low, medium, high) from video data. We explored a set of traditional machine learning algorithms: Support Vector

Machine with polynomial (SVM-poll) and radial basis kernel (SVM-rbf), Random Forest (RF), BayesNet (BN), and Multi-Layer Perceptron (MLP). The experiments were performed using feature reduction and dimensionality reduction techniques:

- PCA - 16 principal components extracted from the data (obtained with a threshold of 95% of variance covered),
- Greedy - 12 features obtained from Greedy Stepwise Search combined with Correlation-based Feature Selection.

We have used a 10-fold validation procedure. The results for the three-class classification task are shown in Table VI. All experiments were performed in Weka 3.8 software[2].

TABLE VI: Results (in terms of F-score) for 3-class classification task

	Engagement		**Confidence**	
	PCA	**Greedy**	**PCA**	**Greedy**
SVM-rbf	0.63	0.63	0.6	0.62
SVM-poll	0.61	0.56	0.56	0.57
RF	0.65	0.73	0.61	0.7
MLP	0.57	0.54	0.49	0.5
BayesNet	0.52	0.6	0.47	0.56

C. Discussion

As can be seen in Table VI, the best results were obtained with Random Forest and Greedy feature reduction. In general, the results are not perfect, but it should be noted that the experiments were performed on noisy 2D data extracted using OpenPose. Some of the tasks given to the children required the rotation of the whole body, leading to a lack of 2D data during the child's rotation. We believe that the results could improve using 3D positional data (e.g., from Kinect[3] or Notch[4]). Another important shortcoming is that we have not used all relevant cues from Tab. III- V. By implementing the remaining cues we hope in future to improve the classification results.

VIII. Conclusions and future steps

In accordance to the the main theme of this year conference which is Affective Computing for ALL (AC4ALL), in this paper, we focused on a specific group of impaired users of AC technology. We hope that this work may help include such specific-needs users in benefiting from AC technologies. In particular, we analyzed full-body indicators of engagement and confidence in visually impaired primary school children. A two-steps annotation was performed by expert and non expert annotators to identify a set of nonverbal cues of engagement and confidence. We also proposed a preliminary classification model of engagement and confidence levels from a set of cues extracted from 2D positional data.

The main contributions of this paper are:

- To the authors' knowledge, this is the first analysis of the nonverbal full-body cues of cognitive states in visually impaired children, and a first attempt to create a model of nonverbal full-body communication for visually impaired children in learning context
- This is one of the first attempts to automatically classify confidence levels from bodily cues in the context of single-user learning tasks

As future steps, we plan to perform a similar analysis of nonverbal behaviors in the same context on a control group of sighted children to further understand differences and commonalities in nonverbal behaviors compared to visually impaired ones. As part of future works, then, we will record more children data, and extend, through other rounds of human annotation, the nonverbal behaviors model. Since the initial results of the automatic classification are promising, we will extract more features to improve the results, as well as test different machine learning techniques. Another possible extension involve multimodal data collection and classification. For this purpose, we consider using low-intrusive wearable sensors that can be used to collect the data of children activity without violating their privacy, such as IMU and EMG sensors.

[2]https://www.cs.waikato.ac.nz/ml/weka/downloading.html

[3]https://developer.microsoft.com/en-us/windows/kinect

[4]https://wearnotch.com

References

[1] Nameera Akhtar and Morton Ann Gernsbacher. On privileging the role of gaze in infant social cognition. *Child development perspectives*, 2(2):59–65, 2008.

[2] Paolo Alborno, Andrea Cera, Stefano Piana, Maurizio Mancini, Radoslaw Niewiadomski, Corrado Canepa, Gualtiero Volpe, and Antonio Camurri. Interactive sonification of movement qualities – a case study on fluidity. In *Proceedings of ISon 2016, 5th Interactive Sonification Workshop*, 2016.

[3] Salvatore M. Anzalone, Sofiane Boucenna, Serena Ivaldi, and Mohamed Chetouani. Evaluating the engagement with social robots. *International Journal of Social Robotics*, 7(4):465–478, Aug 2015.

[4] Cigdem Beyan, Muhammad Shahid, and Vittorio Murino. Investigation of small group social interactions using deep visual activity-based nonverbal features. In *Proceedings of the 26th ACM International Conference on Multimedia*, MM '18, pages 311–319, New York, NY, USA, 2018. ACM.

[5] Ann E Bigelow. The development of joint attention in blind infants. *Development and Psychopathology*, 15(2):259–275, 2003.

[6] Antonio Camurri, Gualtiero Volpe, Stefano Piana, Maurizio Mancini, Radoslaw Niewiadomski, Nicola Ferrari, and Corrado Canepa. The dancer in the eye: Towards a multi-layered computational framework of qualities in movement. In *Proceedings of the 3rd International Symposium on Movement and Computing*, MOCO '16, pages 6:1–6:7, New York, NY, USA, 2016. ACM.

[7] Zhe Cao, Gines Hidalgo, Tomas Simon, Shih-En Wei, and Yaser Sheikh. OpenPose: realtime multi-person 2D pose estimation using Part Affinity Fields. In *arXiv preprint arXiv:1812.08008*, 2018.

[8] Ginevra Castellano, André Pereira, Iolanda Leite, Ana Paiva, and Peter W. McOwan. Detecting user engagement with a robot companion using task and social interaction-based features. In *Proceedings of the 2009 International Conference on Multimodal Interfaces*, ICMI-MLMI '09, pages 119–126, New York, NY, USA, 2009. ACM.

[9] Guillaume Chanel, Cyril Rebetez, Mireille Bétrancourt, and Thierry Pun. Boredom, engagement and anxiety as indicators for adaptation to difficulty in games. In *Proceedings of the 12th International Conference on Entertainment and Media in the Ubiquitous Era*, MindTrek '08, pages 13–17, New York, NY, USA, 2008. ACM.

[10] Arodami Chorianopoulou, Efthymios Tzinis, Elias Iosif, Asimenia Papoulidi, Christina Papailiou, and Alexandros Potamianos. Engagement detection for children with autism spectrum disorder. In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 5055–5059. IEEE, 2017.

[11] Domenic V Cicchetti. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological assessment*, 6(4):284, 1994.

[12] David A Cole and Luanna H Meyer. Social integration and severe disabilities: A longitudinal analysis of child outcomes. *The Journal of Special Education*, 25(3):340–351, 1991.

[13] Peter G Cole and Lee Chee Pheng. The effects of verbal mediation training on the problem-solving skills of children with partial sight and children without visual impairments. *International Journal of Disability, Development and Education*, 45(4):411–422, 1998.

[14] Jacquelynne S Eccles. The development of children ages 6 to 14. *Future of children*, 9(2):30–44, 1999.

[15] Erik Erikson. Youth: Identity and crisis. *New York, NY: WW*, 1968.

[16] Anthony Francescucci and Laila Rohani. Exclusively synchronous online (viri) learning: The impact on student performance and engagement outcomes. *Journal of Marketing Education*, 41(1):60–69, 2019.

[17] Maria Frank, Ghassem Tofighi, Haisong Gu, and Renate Fruchter. Engagement detection in meetings. *CoRR*, abs/1608.08711, 2016.

[18] Jennifer A Fredricks, Phyllis C Blumenfeld, and Alison H Paris. School engagement: Potential of the concept, state of the evidence. *Review of educational research*, 74(1):59–109, 2004.

[19] Bi Ge, Hae Won Park, and Ayanna M. Howard. Identifying engagement from joint kinematics data for robot therapy prompt interventions for children with autism spectrum disorder. In Arvin Agah, John-John Cabibihan, Ayanna M. Howard, Miguel A. Salichs, and Hongsheng He, editors, *Social Robotics*, pages 531–540, Cham, 2016. Springer International Publishing.

[20] JS Gehris, RA Gooze, and RC Whitaker. Teachers' perceptions about children's movement and learning in early childhood education programmes. *Child: care, health and development*, 41(1):122–131, 2015.

[21] Joseph F. Grafsgaard, Joseph B. Wiggins, Kristy Elizabeth Boyer, Eric N. Wiebe, and James C. Lester. Automatically recognizing facial indicators of frustration: A learning-centric analysis. In *Proceedings of the 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, ACII '13, pages 159–165, Washington, DC, USA, 2013. IEEE Computer Society.

[22] E Green, L Stroud, S Bloomfield, et al. Griffiths scales of child development, 2016.

[23] Judith A Hall, Terrence G Horgan, and Nora A Murphy. Nonverbal communication. *Annual review of psychology*, 70:271–294, 2019.

[24] Kevin A Hallgren. Computing inter-rater reliability for observational data: an overview and tutorial. *Tutorials in quantitative methods for psychology*, 8(1):23, 2012.

[25] Susan Harter. The development of self-representations. 1998.

[26] Jana M Iverson and Susan Goldin-Meadow. What's communication got to do with it? gesture in children blind from birth. *Developmental psychology*, 33(3):453, 1997.

[27] Ryosuke Kawamura, Yuushi Toyoda, and Koichiro Niinuma. Engagement estimation based on synchrony of head movements: application to actual e-learning scenarios. In *Proceedings of the 24th International Conference on Intelligent User Interfaces: Companion*. ACM, 2019.

[28] Jaebok Kim, Khiet P. Truong, Vicky Charisi, Cristina Zaga, Vanessa Evers, and Mohamed Chetouani. Multimodal detection of engagement in groups of children using rank learning. In Mohamed Chetouani, Jeffrey Cohn, and Albert Ali Salah, editors, *Human Behavior Understanding*, pages 35–48, Cham, 2016. Springer International Publishing.

[29] Jaebok Kim, Khiet P Truong, and Vanessa Evers. Automatic temporal ranking of childrens engagement levels using multi-modal cues. *Computer Speech & Language*, 50:16–39, 2018.

[30] Michele Capella McDonnall, Brenda S Cavenaugh, and J Martin Giesen. The relationship between parental involvement and mathematics achievement for students with visual impairments. *The Journal of Special Education*, 45(4):204–215, 2012.

[31] RA McWilliam and Amy M Casey. *Engagement of every child in the preschool classroom*. Paul H. Brookes Publishing Company Baltimore, MD, 2008.

[32] AN Meltzoff, R Brooks, R Flom, K Lee, and D Muir. Gaze following: Its development and significance. 2007.

[33] Yukiko I. Nakano and Ryo Ishii. Estimating user's engagement from eye-gaze behaviors in human-agent conversations. In *Proceedings of the 15th International Conference on Intelligent User Interfaces*, IUI '10, pages 139–148, New York, NY, USA, 2010. ACM.

[34] Radoslaw Niewiadomski, Ksenia Kolykhalova, Stefano Piana, Paolo Alborno, Gualtiero Volpe, and Antonio Camurri. Analysis of movement quality in full-body physical activities. *ACM Transaction on Interactive Intelligent Systems*, 9(1):1:1–1:20, February 2019.

[35] Gavin Northey, Tania Bucic, Mathew Chylinski, and Rahul Govind. Increasing student engagement using asynchronous learning. *Journal of Marketing Education*, 37(3):171–180, 2015.

[36] Eda Okur, Nese Alyuz, Sinem Aslan, Utku Genc, Cagri Tanriover, and Asli Arslan Esme. Behavioral engagement detection of students in the wild. In *International Conference on Artificial Intelligence in Education*, pages 250–261. Springer, 2017.

[37] Temitayo Olugbade, Nadia Berthouze, Nicolai Marquardt, and Amanda Williams. Human observer and automatic assessment of movement related self-efficacy in chronic pain: From exercise to functional activity. *IEEE Transactions on Affective Computing*, 2018.

[38] Temitayo Olugbade, Luigi Cuturi, Giulia Cappagli, Erica Volta, Paolo Alborno, Joseph Newbold, Nadia Bianchi-Berthouze, Gabriel Baud-Bovy, Gualtiero Volpe, and Monica Gori. What cognitive and affective states should technology monitor to support learning? In *Proceedings of the 1st ACM SIGCHI International Workshop on Multimodal Interaction for Education*, pages 22–25. ACM, 2017.

[39] Temitayo A Olugbade, Joseph Newbold, Rose Johnson, Erica Volta, Paolo Alborno, Radoslaw Niewiadomski, Max Dillon, Gualtiero Volpe, and Nadia Bianchi-Berthouze. Automatic detection of reflective thinking in mathematical problem solving based on unconstrained bodily exploration. *arXiv preprint arXiv:1812.07941*, 2018.

[40] Claire Cameron Ponitz, Sara E Rimm-Kaufman, Kevin J Grimm, and Timothy W Curby. Kindergarten classroom quality, behavioral engagement, and reading achievement. *School Psychology Review*, 38(1):102, 2009.

[41] H.F.R. Prechtl. Qualitative changes of spontaneous movements in fetus and preterm infant are a marker of neurological dysfunction. *Early Human Development*, 23(3):151 – 158, 1990. New Studies on Movement Assessment in Fetuses and Preterm Infants.

[42] James Rehg, Gregory Abowd, Agata Rozga, Mario Romero, Mark Clements, Stan Sclaroff, Irfan Essa, O Ousley, Yin Li, Chanho Kim, et al. Decoding children's social behavior. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3414–3421, 2013.

[43] C. Rich, B. Ponsler, A. Holroyd, and C. L. Sidner. Recognizing engagement in human-robot interaction. In *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 375–382, March 2010.

[44] Jyotirmay Sanghvi, Ginevra Castellano, Iolanda Leite, André Pereira, Peter W McOwan, and Ana Paiva. Automatic analysis of affective postures and body motion to detect engagement with a game companion. In *Proceedings of the 6th international conference on Human-robot interaction*, pages 305–312. ACM, 2011.

[45] Noor Shaker and Mohammad Shaker. Towards understanding the nonverbal signatures of engagement in super mario bros. In Vania Dimitrova, Tsvi Kuflik, David Chin, Francesco Ricci, Peter Dolog, and Geert-Jan Houben, editors, *User Modeling, Adaptation, and Personalization*, pages 423–434, Cham, 2014. Springer International Publishing.

[46] Hanneke C Sleeuwenhoek, Roeland D Boter, and Adri Vermeer. Perceptual-motor performance and the social development of visually impaired children. *Journal of Visual Impairment & Blindness*, 1995.

[47] David Wechsler. Wechsler intelligence scale for children–fourth edition (wisc-iv). *San Antonio, TX: The Psychological Corporation*, 2003.

[48] Amanda P Williford, Michelle F Maier, Jason T Downer, Robert C Pianta, and Carolee Howes. Understanding how children's engagement and teachers' interactions combine to predict school readiness. *Journal of Applied Developmental Psychology*, 34(6):299–309, 2013.

[49] Alana M Zambone. Serving the young child with visual impairments: An overview of disability impact and intervention needs. *Infants & Young Children*, 2(2):11–23, 1989.