Using fNIRS To Understand Adults' Empathy for Children in AI and Cybersecurity Scenarios

Mohsena Ashraf
University of Colorado Boulder
Boulder, CO, USA
mohsena.ashraf@colorado.edu

Genevieve Patterson
University of Denver
Denver, CO, USA
genevieve.patterson@du.edu

Zachary Kilhoffer
University of Illinois Urbana-Champaign
Champaign, IL, USA
dzk2@illinois.edu

Xu Han
University of Colorado Boulder
Boulder, CO, USA
xu.han-1@colorado.edu

Nolan Brady
University of Colorado Boulder
Boulder, CO, USA
nolan.brady@colorado.edu

Anna Rahn
University of Colorado Boulder
Boulder, CO, USA
anna.rahn@colorado.edu

Nikhitha Atluri
University of Colorado Boulder
Boulder, CO, USA
nich3472@colorado.edu

Violet Oliver
Niwot High School
Niwot, CO, USA
oliver.violet03@svvsd.org

Yun Huang
University of Illinois Urbana-Champaign
Champaign, IL, USA
yunhuang@illinois.edu

Yang Wang
University of Illinois Urbana-Champaign
Champaign, IL, USA
yvw@illinois.edu

Pilyoung Kim
University of Denver
Denver, CO, USA
pilyoung.kim@du.edu

Tom Yeh
University of Colorado Boulder
Boulder, CO, USA
tom.yeh@colorado.edu

ABSTRACT

Empathy for children is critical for designing AI technologies that may affect children. This paper presents the work in progress of a study on the feasibility of a new method to provide objective understanding of people's empathy for children based on functional near infrared spectroscopy (fNIRS). Adult participants (n=13) were presented with benign or concerning scenarios involving children interacting with AI technologies. Their brain activation patterns were recorded and analyzed. Preliminary data analysis revealed distinctive patterns in the mPFC region, which justifies future work to fully realize the potential of this method.

CCS CONCEPTS

• **Human-centered computing**;

KEYWORDS

Empathy for Children, fNIRS, Children-AI interaction, Vignettes

ACM Reference Format:

Mohsena Ashraf, Genevieve Patterson, Zachary Kilhoffer, Xu Han, Nolan Brady, Anna Rahn, Nikhitha Atluri, Violet Oliver, Yun Huang, Yang Wang, Pilyoung Kim, and Tom Yeh. 2023. Using fNIRS To Understand Adults' Empathy for Children in AI and Cybersecurity Scenarios. In *Interaction Design and Children (IDC '23), June 19–23, 2023, Chicago, IL, USA*. ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3585088.3593869

IDC '23, June 19–23, 2023, Chicago, IL, USA

ACM ISBN 979-8-4007-0131-3/23/06.
https://doi.org/10.1145/3585088.3593869

1 INTRODUCTION

As artificial intelligence is growing massively day by day, young children and kids are growing up in a fast forwarding digital world [20]. They are getting more exposed to virtual assistants (e.g., *Google Assistant*, *Alexa*, or *Siri*) and numerous AI based apps, such as— *FaceTime*, *TikTok*, *Whatsapp*, *Youtube*, etc.). These AI based applications have a great potential to help children learn faster, develop their basic knowledge and emotions, and improve their creativity [23, 24]. However, if this exposure is not observed and monitored, that may lead to some fatal consequences and affect their behavioral development and mental growth [8]. Hence, our vision for the future is that children can safely live, grow, and learn in the world of AI as it is becoming ubiquitous. To accomplish this vision, one strategy is to ensure those who are responsible for developing AI technologies can understand, pay attention to, and care about the needs of young children while they design, build, and evaluate those technologies, in other words, having empathy for children.

How can we understand AI developers' empathy toward children? Most existing methods are subjective, such as a survey that asks participants questions like *"do you feel empathetic toward children?"* [2, 12]. This type of questioning is often plagued by subjective biases. In this study, our overall goal is to understand and evaluate if functional near infrared spectroscopy (fNIRS) can be an objective method to assess a person's empathy for children, to complement existing subjective methods. If so, one potential application would be to use this method to validate the effectiveness of design techniques (e.g., participatory design with children) aimed to help developers less experienced with children to become more familiar with and empathetic for children. For this purpose, we plan to study two research questions:

(1) To what extent can fNIRS detect distinctive brain activation patterns when a participant is presented with a *benign* versus a *concerning* scenario involving young children interacting with a piece of technology?
(2) To what extent can those distinctive brain activation patterns be related to empathy for young children?

Next, we review the literature in Section 2. The details for designing the vignettes are described in Section 3 followed by our new fNIRS paradigm in Section 4. Section 5 delineates the behavioral data analysis, and fNIRS data analysis are highlighted in Section 6. We discuss our paper and point out the future work in Sections 7 and 8.

2 RELATED WORK

Given recent advances in technological innovation, interactions between people and technology are ubiquitous in modern society. This is also true for young children, who frequently interact with electronic and screen based media [21, 22]. With the increase in these interactions comes an added risk for concerning interactions between children and technologies that are not designed to be sensitive to their immature cognitive state and at worst are designed to exploit this vulnerability. However, little is known about interaction designers' empathetic responses to scenarios involving young children. Therefore, with this task we aim to investigate patterns of brain activation in response to scenarios describing benign and concerning interactions between young children and technology. By measuring changes in patterns of brain activation in response to this task, we hope to use this task to help additionally quantify the impact of an educational intervention designed to increase understanding of the potential risks to young children in their interactions with unethically designed technologies.

Previous research in social neuroscience has found that similar brain regions are activated when people experience something themselves compared to watching others experience it [5, 13]. These *"shared representations"* of the experiences of the self and others are thought to underlie empathy at a neural level [5, 13]. Tasks designed to elicit empathy have been associated with activation in brain regions including the dorsal anterior cingulate cortex (dACC) and anterior mid cingulate cortex, insula and right temporoparietal junction, ventrolateral prefrontal cortex (vlPFC), orbitofrontal cortex, and mentalizing regions such as the medial prefrontal cortex (mPFC), precuneus, and temporal pole [7]. Empathy related brain activation may also be an important predictor of empathetic behavior. For instance, activation in vlPFC during an empathy eliciting task has been shown to relate to an increase in altruistic behavior (i.e., distributing money to another person) [9]. Therefore, measuring the neural basis of empathetic responding may play an important role in both understanding and predicting prosocial behavior, as well as serve as an indicator of response to interventions designed to increase empathy and related prosocial behavior.

3 DEVELOPING VIGNETTES

A critical component of our method to measure empathy was to develop a collection of scenarios to serve as stimuli to elicit empathetic responses from participants. We began by interviewing 16 educators to understand their views and knowledge regarding cybersecurity and AI ethics education. From these interviews, we identified several cybersecurity and AI ethics issues that the educators believe are important to teach. The issues ranged from inappropriate advertisements, inappropriate contact with strangers via technology, technology addiction, information collection and use by smart toys, microtransactions in video games, interactions with voice agents (e.g., *Alexa*, *Siri*, etc.) and age-inappropriate recommendations provided by them, smart cameras and surveillance, cyberbullying via messenger services, use of online tools for academic cheating, location tracking, and social media apps (e.g., *Tik Tok*).

However, while young children are vulnerable to these risks, most of the available examples are based on real-life events involving teens or adults. Thus, we edited these scenarios to include children as the main subject affected by the issue. To do so, we formed an interdisciplinary team including experts of AI ethics, cybersecurity, developmental psychology, as well as high school students and college freshmen. We conducted a series of writing sessions to brainstorm and write a large set of candidate scenarios. Then, we reviewed these scenarios and identified a subset that were most realistic, probable, harmful, and relatable. For each selected scenario, we developed a *benign* and a *concerning* (ethically fraught) version. For example, one vignette describes a 6-year-old child watching videos on an *iPad* when an advertisement appears on the screen. In the *concerning* version of the scenario the advertisement is for an *R* rated film. In the *benign* version, all details of the scenario are the same except the advertisement that appears is for a *G* rated film. We fine tuned the attributes of the characters to control for confounding variables such as gender, age, and tone. For example, we ensured vignettes contained characters equally representing girls and boys, and ages of 5 – 12. We also ensured grammatical consistency (3rd person perspective), and removed certain gender and age stereotypes (e.g., girls can also do sports and like coding). Finally, we conducted several pilot sessions to identify and fix remaining issues such as lack of clarity and excessive length. In the end, we developed 24 vignettes (12 matched pairs) describing interactions between young children and technology.

4 FNIRS PARADIGM

We designed a new FNIRS paradigm to study people's brain activation patterns when they are presented with scenarios involving children interacting with AI technologies. This paradigm took about 45 minutes and was approved by an Institutional Review Board.

Participants: In our first study using this paradigm, we focused on college students pursing an engineering degree. The rationale is that these students are likely to develop technologies that may affect children in the future. The inclusion criteria are: 18 years or above, fluent in English, and no medical histories of neurological disorders (i.e., epilepsy, migraines, or seizures). Our target sample size is 24. At the time of writing this work-in-progress paper, we collected data from 13 participants.

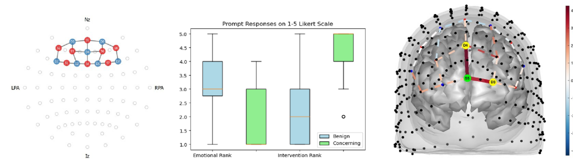
Tools and Equipment: We used the `NIRSx NIRSport2` system with the prefrontal montage consisting of 15 optodes (8 sources and 7 detectors). In this montage setting, eight optode sources ($S1$-$S8$) transmit near infrared light passing through the brain tissues of the prefrontal cortex (PFC) and the seven detectors ($D1$-$D7$) measure the amount of diffusely refracted light, where each optode has an average spatial resolution of 3 cm in between them. Each source and detector pair is known as a channel. Figure 1a depicts the PFC montage structure that has been used in our study. Before starting the data collection, the fNIRS system was calibrated to ensure good signal quality. We also used the open source software *PsychoPy* to display the task stimuli on a laptop screen, and collect the data for behavioral analysis.

fNIRS Setup: Configuring the experimental fNIRS setup for a participant (e.g., Alex) involved four consecutive steps— *1)* First, Alex sat on a comfortable chair in front of a laptop screen placed on a desk, from where he could see the laptop screen easily. *2)* The fNIRS system comes with three cap sizes (52 cm, 56 cm, and 58 cm). After measuring Alex's head size with a tape, a trained researcher placed the appropriate fNIRS cap over their head. The cap was fitted with light sensors, and the wearing experience seemed a little bit tight as a swimming cap at first. *3)* The researcher then adjusted the cap for gaining optimal signal quality without causing any discomfort to Alex as it mostly involved adjusting sensors and removing tension from cables. *4)* Finally, Alex was asked if there is any pain because of the cap. The researcher made additional adjustments to mitigate the pain if Alex felt so.

Vignettes Study: After setting up the fNIRS system, Alex went through the reading of a series of vignettes displayed on PsychoPy, where Alex would be instructed as such: *"Please read the following scenarios and answer the questions that follow"*. The vignettes were presented to Alex in a randomized order, where each one was displayed on the screen for 12 seconds, and was followed by two questions— *i) "What is your emotional reaction to the situation you just read about?"*, *ii) "How motivated would you be to intervene in the situation you just read about?"*. Each question had a duration of 6 seconds, and for each of them, Alex was asked to provide a response using the stimulus presentation of the laptop keyboard. For the first question, Alex had to select an option from a *Likert* scale representing the verbal anchors as such: 1. Very Negative, 2. Negative, 3. Neutral, 4. Positive, 5. Very Positive. For the second question, the verbal anchors were: 1. Not at all, 2. Slightly, 3. Somewhat, 4. Very, and 5. Extremely. After responding to both of the questions, a fixation cross was displayed on the screen for a rest period of 10 seconds. Upon completion of providing all the responses, the researcher carefully removed the cap from Alex's head.

5 BEHAVIORAL DATA ANALYSIS

As the experiment was conducted, *PsychoPy* stored a number of data points regarding the session. The primary dependent variables were the key response logged for the emotional response prompt, the key response for the intervention prompt, and the response times for each keystroke measured in seconds. As was previously mentioned, the key responses were done on a *Likert* scale (1-5). The key response times were all measured in terms of seconds between prompt display and the response being given. Given the prompts, we hypothesized that the concerning prompts would result in a lower reported emotional response and a higher likelihood of intervention.



(a) (b) (c)

Figure 1: (a) A PFC montage structure with 15 optodes. (b) Key responses as recorded from PsychoPy for the emotional and the intervention response. For emotional responses, 1 signifies a negative and 5 indicates a positive emotional response. For intervention, 1 indicates a lower likelihood and 5 indicates a higher likelihood of intervening. (c) fNIRS results on *concerning* vs. *benign* vignettes comparison (front view, *HbO*). Significant channels after FDR correction ($q < 0.05$) are highlighted using solid lines. The figure shows group-level contrast between *concerning* and *benign* vignettes (*benign* > *concerning*, *hotter* colors indicate channels showing a larger positive magnitude contrast between the two conditions).

Benign Prompts: The vignettes were organized into two categories— *concerning* and *benign*. The distinction between the two prompts was done by using similar scenarios with alternative endings. For example, a child using the internet for the sake of finding a tutor would be considered *benign* while the same child using the internet to make unauthorized purchases online would be considered *concerning*. The *benign* prompts had an emotional ranking mean of 3.34 and a response time mean of 2.21 seconds, both with 32 samples and standard deviations of 1.21 and 1.23 seconds respectively. The intervention ranking prompt had a mean of 2.31 and a response time mean of 2.30 seconds, both with 32 samples and standard deviations of 1.31 and 1.51 seconds respectively (see Figure 1b).

Concerning Prompts: We examined 33 sample responses to emotional and intervention ranking prompts. For the emotional ranking response, the mean was 1.72 with a standard deviation of 0.94, and the response time had a mean of 2.17 seconds with a standard deviation of 1.38 seconds. In contrast, the intervention ranking response had a mean of 4.30 with a standard deviation of 1.02, and the response time for the intervention prompt had a mean of 2.26 seconds with a standard deviation of 1.28 seconds.

Results: We used a mixed effects model to assess the interaction between independent variables (prompt type) and dependent variables (emotional ranking, intervention ranking, and response times). Only the prompt type to emotional and intervention rank response interactions were significant. The interaction between prompt type and emotional rank ($n = 65, p < 0.001$) had a coefficient of -1.62 and a 95% confidence interval (CI) of $[-2.14, -1.09]$. The interaction between prompt type and intervention rank ($n = 65, p < 0.001$)

had a coefficient of 0.22 and a 95% CI of [0.154, 0.277]. Results show that the *benign* and *concerning* vignettes had different effects on readers, with the *benign* prompts generating more ambiguity. This was demonstrated by the significant variation in both mean and range of results, as shown in Figure 1b.

6 FNIRS DATA ANALYSIS

Our fNIRS data processing involved three main steps: pre-processing, first-level analysis, and group-level analysis. The analysis was carried out using `MATLAB 2020b` with the `NIRS Toolbox 2022.4.13`. During pre-processing, the raw data was quality checked using a threshold of Scalp Coupling Index (SCI) = 0.8, and NIRS Toolbox automatically handled bad channels and `NaN` values (if any). The raw data was then resampled from 10Hz to 4Hz and converted to *HbO/HbR* concentrations using the modified Beer-Lambert Law (MBLL) [4]. We further conducted first-level analysis to model brain activations for each subject individually. We used the canonical hemodynamic response function (HRF) [18] to run within-subject general linear model (GLM) per subject to model the changes in *HbO/HbR* concentrations under different conditions (i.e., the *concerning* and *benign* vignettes). To control motion artifacts, we fit the GLMs using an AR-IRLS (autoregressive-iteratively reweighted least squares) method [19]. For group-level analysis, we used mixed effects models [10] to estimate averaged regression coefficients (beta values) generated by the first-level analysis. Individual subjects were treated as random effects and the intercept was removed. Post-hoc statistical contrasts (t-test) based on mixed effect objects were performed to determine the mean differences in beta values between conditions. To control for false positives due to multiple comparisons, we applied a false discovery rate (FDR) threshold at $q < 0.05$.

Results: Our t-test analysis of *concerning* vs. *benign* vignettes revealed that the different types of vignettes had a statistically significant effect on HbO brain activation across two channels in the medial prefrontal cortex (`mPFC`, source 5-detector 4 ($S5$-$D4$), source 5-detector 3 ($S5$-$D3$)). Both channels survived FDR correction. Figure 1c shows these results. In mPFC, the *benign* vignettes exhibited higher trending of HbO activation compared to the *concerning* vignettes ($S5$-$D4$, *benign* > *concerning*, $t = 4.27$, $p < 0.001$, $q < 0.05$; $S5$-$D3$, *benign* > *concerning*, $t = 3.48$, $p < 0.01$, $q < 0.05$).

7 DISCUSSION

Results from our task demonstrate that while reading the vignettes participants are activating channels in the mPFC, a region well known for its role in social cognition, mentalizing, and theory of mind [for reviews see [3, 14, 16]]. In particular, we found greater activation in this region for the *benign* rather than *concerning* condition. One potential explanation for this finding is that the *benign* conditions were seen by participants as more ambiguous in the absence of any overtly negative information and they were recruiting mPFC to assist with processing the information in the vignette. This is in line with previous research demonstrating that mPFC plays a role in interpreting ambiguous social information. For instance, one study found that mPFC is activated during processing of ambiguous language [17] and another study utilizing a mental state inference task found that mPFC was more responsive during ambiguous vs unambiguous inferences [11]. Behaviorally, findings from our task suggest that while overall participants were less likely to report a desire to intervene in the *benign* condition, the full range of response options was present in our data with some participants reporting a high likelihood of intervening. This suggests there may have been some degree of ambiguity present. Additionally, while vignettes in the *benign* condition were specifically written to appear neutral, all involved descriptions of young children interacting with technology. There is broader societal debate about young children's interactions with technology in general with previous research finding that parents and teachers in training hold concerns about children's technology use [1, 6, 15]. Therefore, it is possible that even in the absence of overt risk, participants view any interactions between young children and technology as holding some degree of risk, thus influencing the interpretation of the *benign* scenarios as ambiguous.

While the findings are different from what was originally expected, the mPFC activation to the *benign* condition may reflect the mental processes that are involved in empathy specifically in response to more ambiguous situations. The question of how much participants want to intervene will require the participants to think the situations in vignettes from the young child's perspective, and evaluate the vulnerability of the young child. The results of the behavioral ratings demonstrated that most of the participants felt strongly about intervening in response to the concerning condition. The current findings are helpful to guide further studies, and provide the potential that the mPFC activation may be effective in evaluating psychological processes in response to subtle and more ambiguous issues that can occur during interactions with the AI.

8 FUTURE WORK

Future work includes collecting data from more participants, performing comprehensive analysis of the data using the basic empathy scale, and extracting design implications for HCI practitioners. In terms of the FNIRS paradigm, further improvements may involve removing the question on the likelihood of intervening and asking rating questions after each post-scan task, and introducing other conditions such as young children interacting with another child/adult in a benign vs concerning ways, which will allow interaction analysis of agency (AI vs human) and condition (benign vs concerning). Future studies may also consider assessing brain activation in regions beyond PFC as empathy recruits the network of the brain regions in the prefrontal, temporal and parietal lobes.

ACKNOWLEDGMENTS

This work was supported by National Science Foundation under Grant No. 2115004, 2114991, 2115008.

REFERENCES

[1] Cora Bergert, , Antonia Köster, Hanna Krasnova, Ofir Turel, and and. 2020. Missing Out on Life: Parental Perceptions of Children's Mobile Technology Use. , 568–583 pages. https://doi.org/10.30844/wi_2020_f1-bergert

[2] Arnaud Carré, Nicolas Stefaniak, Fanny D'Ambrosio, Leïla Bensalah, and Chrystel Besche-Richard. 2013. The Basic Empathy Scale in Adults (BES-A): Factor structure of a revised form. *Psychological Assessment* 25, 3 (2013), 679–691. https://doi.org/10.1037/a0032297

[3] Sarah J. Carrington and Anthony J. Bailey. 2009. Are there theory of mind regions in the brain? A review of the neuroimaging literature. *Human Brain Mapping* 30, 8 (Aug. 2009), 2313–2335. https://doi.org/10.1002/hbm.20671

[4] Britton Chance, Endla Anday, Shoko Nioka, Shuoming Zhou, Long Hong, Katherine Worden, C. Li, T. Murray, Y. Ovetsky, D. Pidikiti, and R. Thomas. 1998. A novel method for fast imaging of brain function, non-invasively, with light. *Optics Express* 2, 10 (May 1998), 411. https://doi.org/10.1364/oe.2.000411

[5] Frederique de Vignemont and Tania Singer. 2006. The empathic brain: how, when and why? *Trends in Cognitive Sciences* 10, 10 (Oct. 2006), 435–441. https://doi.org/10.1016/j.tics.2006.08.008

[6] Marjory Ebbeck, Hoi Yin Bonnie Yim, Yvonne Chan, and Mandy Goh. 2015. Singaporean Parents' Views of Their Young Children's Access and Use of Technological Devices. *Early Childhood Education Journal* 44, 2 (Feb. 2015), 127–134. https://doi.org/10.1007/s10643-015-0695-4

[7] Haakon G Engen and Tania Singer. 2013. Empathy circuits. *Current Opinion in Neurobiology* 23, 2 (April 2013), 275–282. https://doi.org/10.1016/j.conb.2012.11.003

[8] Eric Greenwald, Maxyn Leitner, and Ning Wang. 2021. Learning Artificial Intelligence: Insights into How Youth Encounter and Build Understanding of AI Concepts. *Proceedings of the AAAI Conference on Artificial Intelligence* 35, 17 (May 2021), 15526–15533. https://doi.org/10.1609/aaai.v35i17.17828

[9] Toshiyuki Himichi and Michio Nomura. 2015. Modulation of empathy in the left ventrolateral prefrontal cortex facilitates altruistic behavior: An fNIRS study. *Journal of Integrative Neuroscience* 14, 02 (June 2015), 207–222. https://doi.org/10.1142/s0219635215500120

[10] Yu Huang, Xinyu Liu, Ryan Krueger, Tyler Santander, Xiaosu Hu, Kevin Leach, and Westley Weimer. 2019. Distilling Neural Representations of Data Structure Manipulation using fMRI and fNIRS. https://doi.org/10.1109/icse.2019.00053

[11] A. C. Jenkins and J. P. Mitchell. 2009. Mentalizing under Uncertainty: Dissociated Neural Responses to Ambiguous and Unambiguous Mental State Inferences. *Cerebral Cortex* 20, 2 (May 2009), 404–410. https://doi.org/10.1093/cercor/bhp109

[12] Darrick Jolliffe and David P. Farrington. 2005. Development and validation of the Basic Empathy Scale. *Journal of Adolescence* 29, 4 (Sept. 2005), 589–611. https://doi.org/10.1016/j.adolescence.2005.08.010

[13] E.J. Lawrence, P. Shaw, V.P. Giampietro, S. Surguladze, M.J. Brammer, and A.S. David. 2006. The role of 'shared representations' in social perception and empathy: An fMRI study. *NeuroImage* 29, 4 (Feb. 2006), 1173–1184. https://doi.org/10.1016/j.neuroimage.2005.09.001

[14] Raymond A. Mar. 2011. The Neural Bases of Social Cognition and Story Comprehension. *Annual Review of Psychology* 62, 1 (Jan. 2011), 103–134. https://doi.org/10.1146/annurev-psych-120709-145406

[15] Pekka Mertala. 2017. Wonder children and victimizing parents – preservice early childhood teachers' beliefs about children and technology at home. *Early Child Development and Care* 189, 3 (May 2017), 392–404. https://doi.org/10.1080/03004430.2017.1324434

[16] Frank Van Overwalle. 2009. Social cognition and the brain: A meta-analysis. *Human Brain Mapping* 30, 3 (March 2009), 829–858. https://doi.org/10.1002/hbm.20547

[17] Joanne L. Powell, Joe Furlong, Christophe E. de Bézenac, Noreen O'Sullivan, and Rhiannon Corcoran. 2019. The Pragmatics of Pragmatic Language and the Curse of Ambiguity: An fMRI Study. *Neuroscience* 418 (Oct. 2019), 96–109. https://doi.org/10.1016/j.neuroscience.2019.08.039

[18] Hendrik Santosa, Xuetong Zhai, Frank Fishburn, and Theodore Huppert. 2018. The NIRS Brain AnalyzIR Toolbox. *Algorithms* 11, 5 (May 2018), 73. https://doi.org/10.3390/a11050073

[19] Hendrik Santosa, Xuetong Zhai, Frank Fishburn, Patrick J. Sparto, and Theodore J. Huppert. 2020. Quantitative comparison of correction techniques for removing systemic physiological signal in functional near-infrared spectroscopy studies. https://doi.org/10.1117/1.nph.7.3.035009

[20] Jiahong Su and Weipeng Yang. 2022. Artificial intelligence in early childhood education: A scoping review. *Computers and Education: Artificial Intelligence* 3 (2022), 100049. https://doi.org/10.1016/j.caeai.2022.100049

[21] George Thomas, Jason A. Bennie, Katrien De Cocker, Oscar Castro, and Stuart J. H. Biddle. 2019. A Descriptive Epidemiology of Screen-Based Devices by Children and Adolescents: a Scoping Review of 130 Surveillance Studies Since 2000. *Child Indicators Research* 13, 3 (July 2019), 935–950. https://doi.org/10.1007/s12187-019-09663-1

[22] Elizabeth A. Vandewater, Victoria J. Rideout, Ellen A. Wartella, Xuan Huang, June H. Lee, and Mi suk Shim. 2007. Digital Childhood: Electronic Media and Technology Use Among Infants, Toddlers, and Preschoolers. *Pediatrics* 119, 5 (May 2007), e1006–e1015. https://doi.org/10.1542/peds.2006-1804

[23] Weiqi Xu and Fan Ouyang. 2022. The application of AI technologies in STEM education: a systematic review from 2011 to 2021. https://doi.org/10.1186/s40594-022-00377-5

[24] Weipeng Yang. 2022. Artificial Intelligence education for young children: Why, what, and how in curriculum design and implementation. *Computers and Education: Artificial Intelligence* 3 (2022), 100061. https://doi.org/10.1016/j.caeai.2022.100061