HUMAN-ROBOT INTERACTION

Robotic reading companions can mitigate oral reading anxiety in children

Lauren L. Wright[1]*, Pooja Vegesna[1], Joseph E. Michaelis[2], Bilge Mutlu[3], Sarah Sebo[1]

Reading fluency is a vital building block for developing literacy, yet the best way to practice fluency—reading aloud—can cause anxiety severe enough to inhibit literacy development in ways that can have an adverse effect on students through adulthood. One promising intervention to mitigate oral reading anxiety is to have children read aloud to a robot. Although observations in prior work have suggested that people likely feel more comfortable in the presence of a robot instead of a human, few studies have empirically demonstrated that people feel less anxious performing in front of a robot compared with a human or used objective physiological indicators to identify decreased anxiety. To investigate whether a robotic reading companion could reduce reading anxiety felt by children, we conducted a within-subjects study where children aged 8 to 11 years ($n = 52$) read aloud to a human and a robot individually while being monitored for physiological responses associated with anxiety. We found that children exhibited fewer physiological indicators of anxiety, specifically vocal jitter and heart rate variability, when reading to the robot compared with reading to a person. This paper provides strong evidence that a robot's presence has an effect on the anxiety a person experiences while doing a task, offering justification for the use of robots in a wide-reaching array of social interactions that may be anxiety inducing.

INTRODUCTION

We are currently in the midst of a global literacy crisis (*1*). Even in developed nations, such as the United States, an analysis of data from the US Department of Education determined that 54% of adults did not score as proficient in literacy (*2*). Research has shown that adult literacy is often determined in childhood—those who fall behind in reading skills during primary school may never catch up (*3*). Teaching interruptions during the COVID-19 pandemic only worsened existing declines in early childhood literacy skills, leaving nearly one-third of elementary school children behind on reading benchmarks (*4*). Literacy rapidly advances during the third through fifth grades when children transition from "learning to read" to "reading to learn," making this an ideal stage of development for interventions to prevent lifelong literacy impairments (*5*). Literacy is highly dependent on reading fluency and reading comprehension, skills that are deeply interdependent in development—to improve reading fluency, one must equally improve comprehension (*6*–*9*). The most common way to practice fluency in the classroom is through reading out loud, typically in front of peers and teachers, which can be a major source of anxiety for some students (*10*, *11*). While reading aloud, an anxious reader's cognitive resources are compromised because they devote cognitive effort to avoiding mistakes, which can come at the expense of comprehension (*12*). In effect, the anxiety experienced while practicing reading fluency has the potential to inhibit reading comprehension and thus literacy development as a whole (*13*).

When a student reads aloud in front of the class, the performance anxiety they may experience can be identified as a form of state anxiety, characterized by a temporary emotional response to a stressful situation rather than an inherent personality trait (*14*). In the case of performance anxiety, the "audience" is the primary stressor (*15*–*17*). For example, one study showed that the composition of the audience was the primary determinant of the anxiety music students experienced, who felt the most anxious in front of their instructors, then their peers, and the least anxious in the presence of a lone recording device (*18*). When considering oral reading anxiety, existing interventions often involve having children read out loud to dolls (*19*), animals (*20*), or a recording device (*21*). Although these interventions may help reduce anxiety, they lack a key benefit of reading aloud to others: identifying when a child mispronounces a word and prompting them to try again. How do we provide instructive feedback during fluency practice without contributing to the burden of performance anxiety?

One promising approach is to use a robot reading companion to alleviate students' performance anxiety while practicing oral reading fluency. Children, particularly 10 to 12 year olds, who read out loud to a robot in the home showed greater reading comprehension and motivation to read at the end of a 2-week study with a robot compared with a worksheet-led control group (*22*). In a study comparing human and robot coreaders with children in a library setting, the robot coreader was seen as a more favorable companion than the human (*23*). Children evaluating humans and robots completing tasks with varying reliability were more biased toward the robot, even an unreliable one, and were more likely to view human mistakes as purposeful (*24*). Although researchers in human-robot interaction (HRI) often try to design robots capable of matching human performance in a given role, these studies provide evidence that there may be specific contexts in which a robot's presence is preferable to a human's. A virtual agent intervention, such as a character on a tablet or a video of a robot, might also be appropriate; however, prior HRI research has found that robots are superior to virtual agents because of the social benefits of physical presence (*25*–*28*). This effect is particularly noted in educational contexts where physically embodied robots outperformed virtual agents in cognitive learning gains (*29*), knowledge recall (*30*), and social presence (*31*).

Given this body of evidence and themes that often arise during interviews and surveys, HRI researchers have speculated that people may feel less performance anxiety doing a task in front of a robot rather than another human. For example, Lin *et al.* (*32*) asked participants to complete puzzles in front of either a human adult or a robot acting as a game guide. Participants under the robot condition were less

[1]University of Chicago, Chicago, IL 60637, USA. [2]University of Illinois Chicago, Chicago, IL 60607, USA. [3]University of Wisconsin-Madison, Madison, WI 53706, USA.
*Corresponding author. Email: llwright@uchicago.edu

likely to agree with the statement "the game guide was judging me." Leyzberg *et al.* (*33*) observed that first-grade students may have been more willing to make mistakes in front of a robot tutor compared with a human one. Michaelis and Mutlu (*22*) found that 10- to 12-year-old children felt less judged by the robot during oral reading than they did in their general experience with their peers, teachers, or parents. However, few studies to date have empirically tested whether people feel less performance anxiety in front of a robot compared with a human.

In this work, we investigated whether the social presence of a robot lessens the anxiety experienced by children practicing oral reading fluency compared with the presence of a human. We conducted a within-subjects user study where 52 children ages 8 to 11 years read out loud individually to both a human and a robot in a counterbalanced order while we monitored objective physiological indicators of anxiety: vocal jitter (*34*), heart rate variability (HRV) (*35*), and facial temperature (*36*) (see Fig. 1 and Movie 1). We anticipated that children would exhibit fewer bodily responses associated with anxiety while reading to the robot versus the human. If robots do cause less anxiety for children, then robots could potentially be used as reading assistants to benefit children who struggle to develop literacy skills because of anxiety.

RESULTS

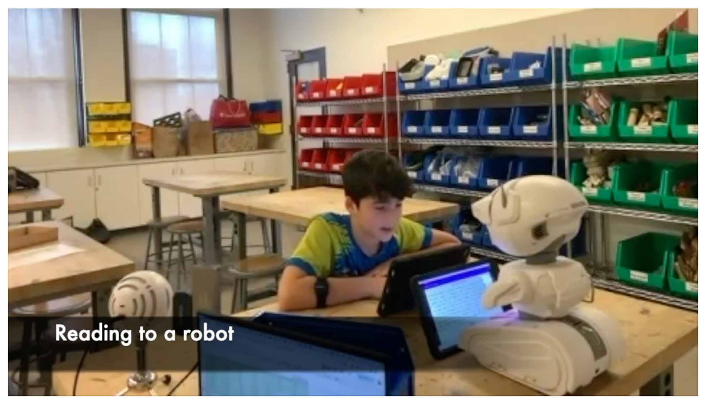
We recruited children ages 8 to 11 [$n = 52$, mean (M) = 9.52 years, and SD = 0.96 year] to participate in a study where they read one passage aloud alone (to establish a physiological baseline) and then two passages each to a robot and to a human, the order of which was counterbalanced across participants. The passages, a mixture of fiction and nonfiction, were taken from the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) 8th Edition oral reading fluency materials (*37*) and took ~5 min each to read. While reading, children received periodic instructive feedback by being asked to repeat three words after completing a page—either preidentified vocabulary words (such as "seltzer," "extraterrestrial," and "aristocratic") or words that either the human or an algorithm identified as being mispronounced. Children read the passages and repeated words at their own pace under both conditions. In between reading with the human and the robot, children completed a small puzzle as an anxiety-reducing exercise (*38*) to bring their bodies back to their physiological baseline. Children also completed short comprehension tests after the passages they read aloud to both the human and the robot. At the end of the study, we conducted a short semistructured interview (*39*).

Physiological responses

Human bodies have involuntary physiological responses to stress. Anxiety is often accompanied by increased activity in the sympathetic nervous system (SNS), triggering our "fight or flight" response (*40*). Increased SNS activity affects a variety of bodily systems, some of which can be measured noninvasively, making it a convenient and objective indicator for anxiety. Because all bodies are unique, measures were taken relative to each individual's baseline, which was established while they read out loud alone. Here, we used voice characteristics, heart rates, and facial temperature changes to determine the relative degree of anxiety the children experienced while reading aloud to a robot and to a human.

Voice characteristics

Anxiety audibly manifests itself through voice instability, wherein an increased SNS response leads to variations in vocal fold vibrations, causing shakiness (*34*). This shakiness results in higher jitter, a measure of cycle-to-cycle variations in the fundamental frequency of a voice. Jitter is calculated by extracting the fundamental frequency from raw audio and then computing the average absolute



Movie 1. Study protocol. Footage from the study demonstrating the study protocol, where a participating child read out loud individually to a robot and to a human while we monitored them for physiological indicators of anxiety.

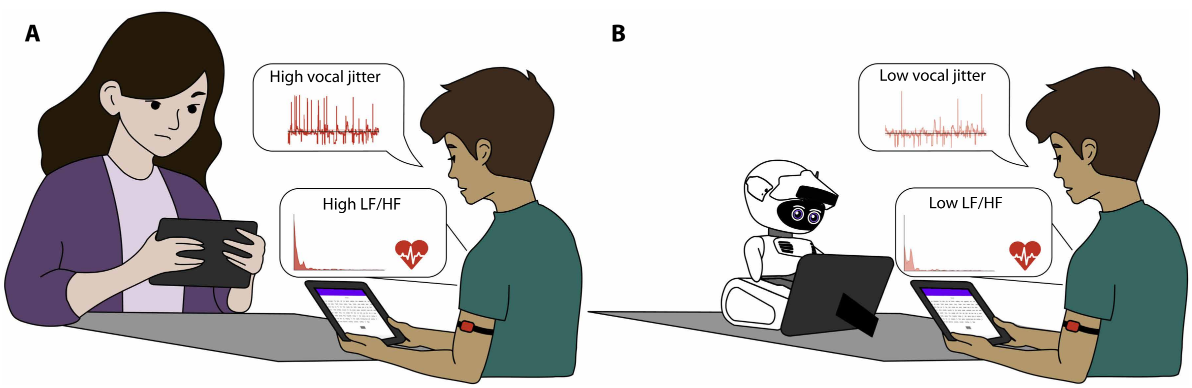


Fig. 1. Study setup. (**A**) Children reading aloud to a human show higher levels of anxiety through higher vocal jitter and a higher LF/HF ratio. (**B**) Children reading aloud to a robot show lower levels of anxiety through lower vocal jitter and a lower LF/HF ratio.

differences between successive periods divided by the average period of pitch, expressed as a percentage of pitch (*34*). Jitter can be calculated from raw audio signals, making it an ideal measure in field experiments. Figure 2A shows the fundamental frequency extracted from 2.5-s audio clips from the same child (P47) reading to both the human and the robot. While reading with the human, this child's fundamental frequency had numerous large deviations. In contrast, their fundamental frequency while reading with the robot had fewer deviations, and most of the deviations were smaller, resulting in lower jitter.

To account for individual vocal characteristics, we subtracted each child's baseline average jitter from their average jitter under each condition to compare the change in jitter between conditions, as shown in Fig. 2B. On average, we found that children's voices reading with the robot had less jitter than their baseline, with an average change of −0.021 percentage point (SD = 0.085), which was significantly lower than the 0.148 percentage point increase in jitter heard when reading with a human [SD = 0.039, $F\,(1,1) = 41.65$, $P < 0.0001$, and $\eta_p^2 = 0.3$]. Children reading out loud to the robot had more stable voices with less jitter than when they read out loud to the human, indicating that, overall, they experienced less anxiety with the robot.

Heart rate

Increased SNS activity because of anxiety is often accompanied by a decrease in HRV, a measure of the variation in time intervals between consecutive heartbeats (*35*), which results in a more regular heart rate. One way to measure HRV is to decompose the changes in heart rate over time into low- and high-frequency bands using the power spectral density (PSD) and then assess the relative power of those bands. Low-frequency (LF) power is associated with increased SNS activity, whereas high-frequency (HF) power is not, which means that anxiety can manifest as an increased LF/HF ratio (*41*, *42*). As anxiety increases, LF begins to dominate as the intervals between heartbeats become more regular (decreasing HRV), which can be measured by analyzing the changes over time from a simple heart rate monitor. Although LF/HF holds some controversy as a measure of HRV during long-term (24 hours) active recordings, during short-term recordings on the order of 10 min focusing on a psychologically related SNS activity, it is still applicable (*43*) and is consistently used as a measure of psychological stress in both adults and children (*44*–*50*). In Fig. 3A, we display the PSD plots for the same child (P14) reading with the human and the robot. In the PSD plot where the child was reading with the human, the area under the curve in the LF band (0.04 to 0.15 Hz) was relatively large, and the area under the curve in the HF band (0.15 to 0.4 Hz) was smaller, leading to a higher LF/HF. While reading with the robot, the area in the LF band was smaller, and the area in the HF band was mostly unchanged, leading to a lower LF/HF.

We compared the LF/HF ratio for every child under each condition with their individual baseline LF/HF ratio. Reading with either the human or the robot resulted in children displaying, on average, an increased LF/HF ratio from their baseline. However, reading with the robot resulted in a significantly smaller increase ($M = 0.951$ and SD = 0.624) than reading with the human ($M = 1.64$, SD = 0.712, $F = 5.96$, $P = 0.016$, and $\eta_p^2 = 0.06$) as seen in Fig. 3B. The HRV of children reading to the robot was composed of lower LF power and higher HF power, indicating that they experienced less anxiety than while reading aloud to the human.

Facial temperature

The SNS activity triggered by anxiety can also alter blood flow, leading to temperature changes in the face (*36*). Thermal imaging is a

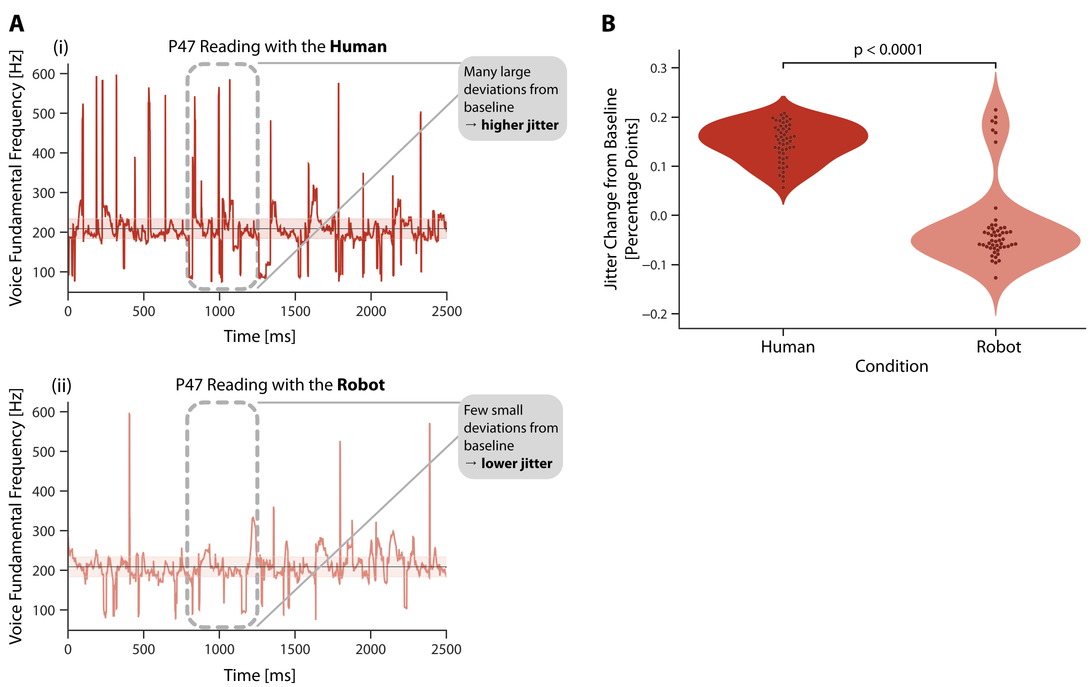


Fig. 2. Voice changes. (**A**) Fluctuations in the fundamental frequency (F0) from 2.5-s audio segments of the same child (P47) reading with the human (i) and the robot (ii) compared against the band of typical fluctuation around P47's baseline average F0. While reading with the human, many large fluctuations illustrate higher jitter, whereas fewer small fluctuations while reading with the robot illustrate lower jitter. (**B**) We observed a significant difference between conditions ($P < 0.0001$) in the average change in children's vocal jitter from baseline values, indicating that they experienced more anxiety reading to the human than to the robot.

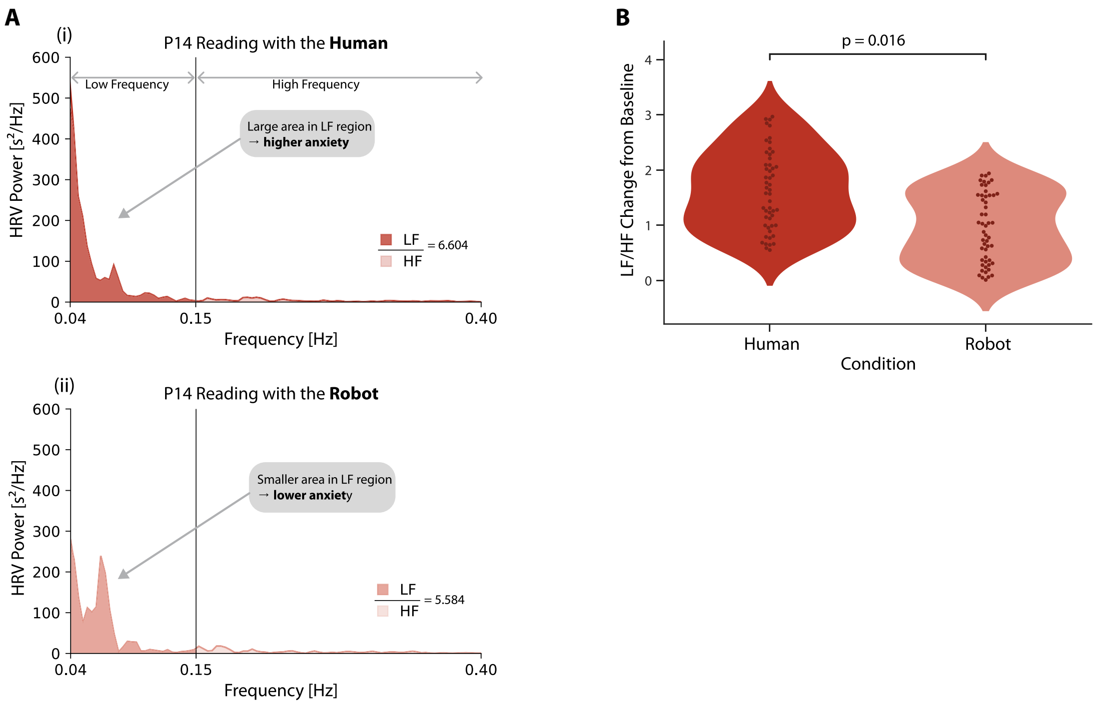


Fig. 3. Heart rate changes. (**A**) PSD plots of the HRV of the same child (P14) reading with the human (i) or the robot (ii) with LF and HF ranges demarcated. While reading with the human, the larger LF/HF ratio was driven by a larger LF power (in the 0.04- to 0.15-Hz range). While reading with the robot, the LF power was not as high, leading to a smaller LF/HF ratio. (**B**) We observed a significant difference between conditions ($P = 0.016$) in the average change in the children's LF/HF ratios from baseline values, indicating that they experienced more anxiety reading to the human than to the robot.

noninvasive way to assess physiological responses to emotional states, such as anxiety. Temperature increases in the periorbital region, the area surrounding the eye, are positively correlated with anxiety (*51*). Children's facial temperatures were measured during the study using a low-cost thermal camera with a 256 pixel–by–192 pixel infrared resolution, mounted at a three-quarter angle toward their faces to prevent the camera from intruding between the child and the reading partner. This meant that only the right side periorbital region was fully visible, from which the average temperature in the periorbital region was extracted [given a reasonable expectation of facial symmetry (*52*)].

Because this measure required a direct view of the orbital region, children with glasses or hair that persistently occluded the face were removed from the facial temperature analysis, leaving $n = 35$ children. We did not observe a significant difference in the changes in temperature in the periorbital region from the baseline when children read to the robot [$M = -0.109$°F (-0.0606°C) and SD = 0.89°F (0.49°C)] compared with when they read to the human [$M = 0.03$°F (-0.016°C), SD = 0.87°F (0.48°C), $F(1,1) = 1.54$, $P = 0.219$, and $\eta_p^2 = 0.02$]. However, in addition to the reduced sample size, this result may also have been affected by environmental variables outside of the control of the researchers, such as ambient room temperature and participants' activity level directly before the study, which may have led to heightened overall body temperature, obscuring any changes that could be observed during the study.

Learning responses

After completing each reading session, children were asked to complete a five-question multiple-choice comprehension test for each passage. Although we had expected to see that children who read with the robot improved their reading comprehension, we did not see a significant difference in the average comprehension score (of 10) between reading sessions with the robot ($M = 6.50$ and SD = 2.12) versus the human [$M = 6.14$, SD = 2.02, $F(1,1) = 2.62$, $\eta_p^2 = 0.03$, and $P = 0.108$]. Although this does not necessarily support our assumption that the decreased anxiety experienced while reading aloud to the robot would help increase reading comprehension, we did not see a decrease in reading comprehension either. One of the arguments often made against robots in educational settings is that they are more distracting than helpful. Here, we show a lack of any substantial interaction between the reading partner and the comprehension scores, which could indicate that the anxiety-mitigating effects of reading to the robot did not come at the expense of comprehension.

Perceptions and preferences

After the study concluded, participating children were invited to share their thoughts on the experiment in a short semistructured interview, asking which form of reading they preferred (alone, with the robot, or with the human) and why. Children who chose to express a preference almost unilaterally chose the robot as their preferred reading partner ($n = 21$ among the 27 who expressed a preference; the remaining 25 participants declined to answer, either electing not to choose or being indifferent to the question). Although children's justifications were brief, we conducted a thematic analysis of the 27 responses to supplement our physiological findings and found participant justifications to converge into two themes: Children had divergent views of the appeal of the robot, and children who preferred reading with the robot did so because of a greater feeling of comfort.

Some children preferred the robot because they found it more appealing, stating "it's really cute" (P16), "I thought it was cool" (P19), and "it felt fun" (P28). However, this was not universal; of the six children who preferred to read either alone or with the human ($n = 4$ preferred alone and $n = 2$ preferred the human), three mentioned discomfort with the robot: "It creeped me out" (P7), "the voice is scary" (P14), and "it's a mechanical thing, so I wasn't as comfortable" (P31). Notably, only one child who chose the human did so seemingly because of a desire for a personal audience: "I like performing" (P15).

Children seemed to feel more comfortable reading in front of the robot than another person. Children stated that, with the robot, they did not feel judged: "The robot was less stressful… the robot is easier because you feel less judged because robots don't have feelings" (P6), and "Even when I made a mistake, I knew it couldn't be mad at me" (P38). Children believed that the robot was there primarily to help: "It just wants to get me to read better" (P44), and "It felt like she was working with me" (P25). They preferred the robot because they did not like reading to people: "I usually mess up more when I'm reading out loud with someone" (P50), and "I was definitely more nervous with the human" (P40). One child summed up their preference by indicating that the robot provided a lower-stress audience while still helping the student: "If I'm reading alone, I don't know if I'm reading something the right way, but if I'm reading with a person, I'm scared I'm not reading something the right way. When I'm reading with the robot, it'll just point it out because it's a robot" (P13). Overall, although children's perceptions of the appeal of the robot differed, most preferred reading with a robot and found it to be helpful to their reading.

DISCUSSION

Researchers in HRI have long suspected that people feel less anxious, less judged, and more comfortable when doing a task in the presence of a robot compared with a human. However, these suspicions were built primarily through observation (*22*, *33*) and subjective measures (*32*). In this work, we contributed objective evidence to support this long-held assumption by demonstrating that children did indeed display fewer physiological indicators for anxiety when reading out loud to a robot compared with a human. In addition, the anxiety-mitigating effect of reading with the robot did not come at the expense of reading comprehension. Our results provide substantial justification for the use of robots not only in education but also in other contexts where a social interaction may induce state anxiety with negative consequences.

The use of objective measures to assess anxiety

Subjective measures used to assess internal states, such as surveys, can be sensitive to each individual's internal biases; for example, answering according to what they believe the researcher wants (*53*), conscious refusal to identify with a "bad" label (*54*), or subconsciously skewing an answer to maintain self-image (*55*). This sensitivity to bias is more pronounced in children, who rarely have the introspective ability needed to name and identify their own emotions, let alone accurately report them (*56*). In this study, we primarily relied on objective, measurable physiological responses with well-documented causal links between nervous system arousal and internal states to assess whether children were likely experiencing anxiety. Even consciously attempting to suppress anxiety does not eliminate the SNS response (*57*), giving us a high degree of confidence in the conclusions made from our results. Objective measures such as the ones used in this study can be a powerful supplement to subjective measures, adding nuance and confidence to conclusions drawn about participants' internal states.

Two of our measures, vocal jitter and HRV, were significantly different when children read with the robot compared with when they read with the human. Children reading with the human had shakier voices and more regular heart rates, indicating that their SNSs were activated, a response to stress that is correlated with anxiety. In contrast, while reading with the robot, children's voices were steadier than their baseline, and their heart rates maintained more variability, which indicates that their nervous systems were not reacting strongly to any perceived threat. In short, our results indicate that children did indeed experience more anxiety while reading to the human than while reading to the robot.

Although our facial temperature results do not align with our voice and heart rate results, this may be because of outside factors rather than a true lack of reaction. Participating schools rarely had space to conduct the study close to the third- to fifth-grade classrooms. Often, escorting a participating child from their class to the study room involved a walk including several flights of stairs, elevating their body temperature. Sometimes, children were participating in the study directly after recess or physical education classes. Although activity-related heart rate changes normalize quickly, temperature changes do not always normalize quickly after physical activity (*58*). To minimize their time out of class, we did not wait for children's body temperatures to normalize before beginning the study. These factors could have been eclipsing any changes due to anxiety. However, we still identify the potential for this measure as a useful, low-cost, and noninvasive method for other researchers to assess internal emotional states in human participant research, albeit best used in more controllable study environments.

Potential learning outcomes

We did not find that the reading companion had any measurable effect on reading comprehension. Although we identified a potential to see reading comprehension gains in this single-session study, we recognize that reading comprehension improvement is more complex to assess than physiological responses and that differences in reading comprehension may only be observed over longer periods of time. Although we anticipated that decreased anxiety could potentially boost reading comprehension, many more factors may have influenced reading comprehension, which could have obscured the effect of anxiety, such as difficulty of the passages, age, or session number. Our comprehension assessments were kept brief to minimize time taken out of class, but it is possible that a difference in comprehension scores could have been seen after children participated in multiple sessions with the robot over a long-term deployment with more extensive assessments. Despite these limitations in assessing comprehension, our single-session study still successfully demonstrated a difference in how children responded physiologically, which is a result we expect to hold up in a long-term deployment that may further investigate comprehension gains.

One of the benefits of reading with the robot rather than alone or to an inanimate object is the potential to provide corrective feedback when the child makes mistakes. To maintain a simplified procedure, in our study, both the human and the robot delivered corrective feedback by asking children to repeat three words at the end of each page they read. The three words were either selected from the list of words that the reading companion had identified as mispronounced (hand selected by the human or algorithmically determined through speech recognition by the robot) or, in the absence of reading errors, preselected vocabulary words. Children under both reading conditions were not told the correct pronunciations, merely asked to repeat the highlighted words. This method of corrective feedback, a form of repeated readings (*59*), represents the bare minimum of what technology can provide but pales in comparison with what a human tutor could offer. Children participating in our study may not have immediately benefited from the feedback we offered here, but because technology advances, the potential for more personalized tutoring with a robot becomes possible.

Implications for the future

Learning, for many people, is often an exercise in vulnerability. Our self-image and confidence can be challenged because we confront the limits of our knowledge and ability. Although this study focused on performance anxiety while reading aloud, education is rife with situations that may induce anxiety. Reducing or eliminating anxiety may not carry a student effortlessly through their curriculum, but it could potentially decrease the likelihood that a student's mental state is the sole determinant of their academic achievement. Robots offer a possible solution—deployed strategically in education, social robots could be used as interventions for students who are academically hindered by their anxiety. Robots could provide a safe space to learn, aiding in providing equal opportunities for students to succeed.

Education is not the only context in which a robot could help mitigate anxiety. Any social interaction that requires vulnerability from one party could be a potential use case for a social robot. Robots used to assist doctors might allow patients to be more honest about pain and discomfort that they feel or about other medications they may be taking. Even physical therapy and rehabilitation exercises, which can sometimes be embarrassing and difficult for patients to perform correctly, might be more tolerable when done while supervised by a robot at the direction of a physical therapist. Any type of skill acquisition that could benefit from feedback, such as singing, painting, or weight lifting, might inspire less hesitance from novice learners when a robot is delivering the help. Our study helps motivate the need to investigate contexts where social robots could be particularly helpful in reducing anxiety.

MATERIALS AND METHODS

We chose a within-subjects experimental design to explore whether the social presence of a robot reduces the anxiety experienced by children practicing oral reading fluency with a human. To assess children's anxiety, we relied on objective physiological measures: vocal jitter, HRV, and facial temperature. We had children read aloud to themselves to establish their physiological baseline, and then children read passages to either a robot or a human, receiving corrections after every page of the passages, after which they read to the other reading companion. Few studies have used these physiological measures in experimental settings with robots, so we chose to conduct a one-time isolated study as an initial attempt to investigate the anxiety felt by children reading aloud to a robot and a human before conducting more naturalistic long-term studies. In addition to anxiety, we also investigated the potential effect of the reading companion (human or robot) on children's comprehension of the reading material.

Study protocol

Participating children were escorted from their classroom to the study room, where they read five texts taken from the DIBELS 8th Edition oral reading fluency passages (*37*). The first passage was read out loud by themselves to establish a physiological baseline, followed by two sessions where they read two passages to an autonomous Misty Robot (*60*) and two passages to one of two researchers (both women aged 18 to 30 years old), counterbalanced in order. After completing the first session with either the human or robot, each child completed a small puzzle to allow their body to return to baseline. Reading passages were administered to the child via a dedicated tablet application that we developed. A paired tablet was either held by the human or propped up in front of the robot. After each page in each story read to a companion, children would indicate that they were finished, at which point they were asked to repeat three words highlighted on the same page—either preidentified vocabulary words (for example, seltzer, extraterrestrial, or aristocratic) or words that either the human or an algorithm [leveraging Double Metaphone phonetic encoding (*61*) with the Levenshtein edit difference (*62*); see the Supplementary Materials for details] identified as being mispronounced. No other help or feedback was provided beyond instructing the child to repeat words under either condition. Children were not prompted to hurry or adjust their reading speed by either reading companion, reading each page and repeating words at their own pace. After reading the two stories under each condition, the child completed five comprehension questions per story. At the end of the study, children were asked whether they had a preference for reading alone, with the human, or with the robot and were invited to ask questions about the study. They were given a piece of candy in thanks for participating before being escorted back to their classroom.

Data collection

An a priori power analysis was conducted using G*Power, assuming a repeated-measures analysis of variance (ANOVA), within factors. A sample size of $n = 52$ was suggested, using a large effect size and a power of 0.80 ($f = 0.2$, $\alpha = 0.05$, $n_{\text{groups}} = 2$, $n_{\text{measurements}} = 2$, correlation = 0.5, and $\epsilon = 1$). We recruited participants, aged 8 to 11 ($n = 52$, $M = 9.52$ years, and SD = 0.96 years), from third- through fifth-grade classes at two different partner schools.

For this study, we obtained both parental consent and child assent before collecting any data. We recorded video footage of each participant over the entirety of the study. During the study, we collected both physiological data, using dedicated sensors monitoring throughout the study, and reading comprehension data, collected after children read with each individual companion. At the end of the study, we also led a semistructured interview. This study was approved by the University of Chicago Institutional Review Board under IRB23-0857.

Physiological measures

Each physiological measure had a dedicated, noninvasive sensor. These sensors were pointed out and explained to participating children during the child assent procedure. Every parent consented and child assented to the use of these sensors.

Learning measures

A five-question multiple-choice comprehension test was given for each passage read with a reading companion, where each question had four answers to choose from. Children were informed that their score on the comprehension tests was for the study only and would not be shared with their parents or teachers. The comprehension tests were created for this study, and questions were focused on detail recall rather than analysis of themes present in the passages. Comprehension scores were averaged over the 10 questions per condition.

Interview
At the end of the study, we conducted a semistructured interview where we asked children whether they had a preference for reading alone, to the robot, or to the human. If they elected to answer, then we asked them why they chose that preference. If children did not elect to answer, then we did not pressure them to choose a preference.

Data analysis
Physiological measures
The raw physiological data first required processing to calculate the metrics of interest. To analyze changes in children's voices, raw audio files, manually truncated for each reading session, were first processed using Adobe Premiere to attenuate background noise. Each segmented and cleaned audio file was then analyzed with Praat (*63*), an audio processing software, through a Python interface (*64*) to calculate the jitter values as the average absolute differences between successive periods divided by the average period of pitch.

We used LF/HF to assess anxiety expressed through HRV. Although LF/HF has some controversy as a measurement of physical stress because of multiple body systems influencing LF power, the relationship between LF power and SNS activity due to psychological stress remains sound (*43*). Less-controversial metrics such as respiratory sinus arrhythmia require the measurement of respiration in addition to heart rate. This would necessitate use of an electrocardiogram (ECG) or a high-resolution photoplethysmogram (PPG), both of which are more invasive and distracting sensors, requiring the application of electrodes on bare skin (ECG), which could itself introduce anxiety during the beginning of the study (*65*, *66*), or wearing a finger clip (PPG), which could be distracting while using a touchscreen tablet. Therefore, we chose LF/HF to assess anxiety because it is less invasive, less distracting, and still consistently used as a measure for psychological stress in both adults (*44*–*46*) and younger populations (*47*–*50*). We used 0.04 to 0.15 Hz as our LF range and 0.15 to 0.4 Hz as our HF range in line with several pediatric studies in our age range (*67*–*71*). The raw heart rate recordings were logged at a sampling rate of 1 Hz, from which we calculated HRV. We computed the PSD using Welch's method (*72*) on the time series of heart rate changes. From the PSD, the areas within the LF and HF ranges were approximated using the trapezoid method. LF/HF was taken as the ratio between these areas.

Thermal video recordings were manually truncated for each reading session. We used FFmpeg (*73*) to take screenshots at a rate of 1 frame/s from the segmented videos to be used for facial recognition. We used the yolov5l_face model (*74*) to identify faces to extract the average temperature from the periorbital region. We then calculated the temperature for every pixel within the periorbital region. To exclude the temperatures of cooler areas such as eyelashes and eyebrows, we calculated the average periorbital area temperature as the average of the 10% warmest pixels within the box (*75*).

Statistical analysis
To analyze the effect of a robot's versus a human's presence on the anxiety experienced by children reading aloud, we used two-way repeated-measures ANOVA tests with the experimental condition and session number as fixed factors and age as a covariate. We ensured that our data held to ANOVA assumptions by verifying that our data were normally distributed and had homogeneity of variance. We also checked for outliers, which were only present in the data for jitter under the robot condition (seven outliers, as can be confirmed visually in Fig. 2A). We ran the ANOVA for jitter both with and without the outliers included, and the significance of the result was not affected by the presence of outliers. Therefore, we chose to include the jitter outliers in the final analysis to maintain transparency. We report the effect size as partial eta squared (η_p^2).

Thematic analysis
All interviews were video recorded, transcribed verbatim, and coded using a reflexive thematic analysis (*76*) approach. Following this method, we began with an initial round of open semantic-level coding of the interview transcripts to capture the content of each turn of talk. We then conducted a round of focused axial coding of the initial codes to identify and generate candidate themes across and within the data. These candidate themes were iteratively refined and discussed by the first and last author and included input from all authors. The final theme was determined by consensus and is supported by direct quotation from study participants.

Supplementary Materials
The PDF file includes:
Materials and Methods
Figs. S1 to S3

Other Supplementary Material for this manuscript includes the following:
MDAR Reproducibility Checklist

REFERENCES AND NOTES

1. S. P. Katiyar, "Adult illiteracy: A global social problem" in *The Palgrave Handbook of Global Social Problems*, R. Baikady, S. M. Sajid, J. Przeperski, V. Nadesan, M. Rezaul Islam, J. Gao, Eds. (Palgrave Macmillan, 2021), pp. 1–20.
2. J. Rothwell, "Assessing the economic gains of eradicating illiteracy nationally and regionally in the United States" (Gallup, Inc., 2020), pp. 3–9; www.barbarabush.org/wp-content/uploads/2020/09/BBFoundation_GainsFromEradicatingIlliteracy_9_8.pdf.
3. E. Ferrer, B. A. Shaywitz, J. M. Holahan, S. E. Shaywitz, Early reading at first grade predicts adult reading at age 42 in typical and dyslexic readers. *NPJ Sci. Learn.* **8**, 51 (2023).
4. Amplify, "Amid academic recovery in classrooms nationwide, risks remain for youngest students with least instructional time during critical early years" (2020); https://amplify.com/pdf/uploads/2022/02/mCLASS_MOY-Results_February-2022-Report.pdf.
5. S. B. Zakariya, "Learning to read, reading to learn: Why third-grade is a pivotal year for mastering literacy" (Center for Public Education, 2015), pp. 1–15; https://cdn-files.nsba.org/s3fs-public/reports/NSBA_CPE_Early_Literacy_Layout_2015.pdf.
6. K. DiSalle, T. Rasinski, Impact of short-term intense fluency instruction on students' reading achievement: A classroom-based, teacher-initiated research study. *J. Teach. Action Res.* **3**, 1–13 (2017).
7. Y.-S. Kim, Y. Petscher, C. Schatschneider, B. Foorman, Does growth rate in oral reading fluency matter in predicting reading comprehension achievement? *J. Educ. Psychol.* **102**, 652–667 (2010).
8. S. J. Samuels, The method of repeated readings. *Read. Teach.* **32**, 403–408 (1979).
9. D. LaBerge, S. J. Samuels, Toward a theory of automatic information processing in reading. *Cog. Psych.* **6**, 293–323 (1974).
10. G. Ramirez, L. Fries, E. Gunderson, M. W. Schaeffer, E. A. Maloney, S. L. Beilock, S. C. Levine, Reading anxiety: An early affective impediment to children's success in reading. *J. Cogn. Dev.* **20**, 15–34 (2019).
11. E. M. Phillips, Anxiety and oral competence: Classroom dilemma. *Fr. Rev.* **65**, 1–14 (1991).
12. S. Zaccoletti, G. Altoè, L. Mason, The interplay of reading-related emotions and updating in reading comprehension performance. *Br. J. Educ. Psychol.* **90**, 663–682 (2020).
13. J. A. Tysinger, P. D. Tysinger, T. D. Diamanduros, The effect of anxiety on the measurement of reading fluency and comprehension. *Educ. Res.* **8**, 2 (2010).
14. T. J. Huberty, Test and performance anxiety. *Prin. Lead.* **10**, 12–16 (2009).
15. D. L. Hamann, An assessment of anxiety in instrumental and vocal performances. *J. Res. Music Educ.* **30**, 77–90 (1982).
16. M. Fredrikson, R. Gunnarsson, Psychobiology of stage fright: The effect of public performance on neuroendocrine, cardiovascular and subjective reactions. *Biol. Psychol.* **33**, 51–61 (1992).
17. M. G. Craske, K. D. Craig, Musical performance anxiety: The three-systems model and self-efficacy theory. *Behav. Res. Ther.* **22**, 267–280 (1984).
18. A. LeBlanc, Y. C. Jin, M. Obert, C. Siivola, Effect of audience on music performance anxiety. *J. Res. Music Educ.* **45**, 480–496 (1997).

19. R. N. Bogatay, "Reading to dolls: Impacts on self-efficacy and oral reading fluency," thesis, Chicago School of Professional Psychology, Chicago, IL (2022).
20. S. S. Hall, N. R. Gee, D. S. Mills, Children reading to dogs: A systematic review of the literature. *PLOS ONE* **11**, e0149759 (2016).
21. C. Brown, "The impact of having second grade students record a text read out loud in relation to their fluency skills," thesis, University of Wisconsin–La Crosse, La Crosse, WI (2020).
22. J. E. Michaelis, B. Mutlu, Reading socially: Transforming the in-home reading experience with a learning-companion robot. *Sci. Robot.* **3**, eaat5999 (2018).
23. H.-P. Yueh, W. Lin, S.-C. Wang, L.-C. Fu, Reading with robot and human companions in library literacy activities: A comparison study. *Br. J. Educ. Technol.* **51**, 1884–1900 (2020).
24. R. Stower, A. Kappas, K. Sommer, When is it right for a robot to be wrong? Children trust a robot over a human in a selective trust task. *Comput. Hum. Behav.* **157**, 108229 (2024).
25. J. Li, The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots and virtual agents. *Int. J. Hum. Comput. Stud.* **77**, 23–37 (2015).
26. E. Deng, B. Mutlu, M. J. Mataric, Embodiment in socially interactive robots. *Found. Trends Robot.* **7**, 251–356 (2019).
27. J. M. Kraus, F. Nothdurft, P. Hock, D. Scholz, W. Minker, M. Baumann, "Human after all: Effects of mere presence and social interaction of a humanoid robot as a co-driver in automated driving" in *AutomotiveUI '16 Adjunct: Adjunct Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Association for Computing Machinery, 2016), pp. 129–134.
28. W. A. Bainbridge, J. Hart, E. S. Kim, B. Scassellati, "The effect of presence on human-robot interaction" in *RO-MAN 2008-The 17th IEEE International Symposium on Robot and Human Interactive Communication* (IEEE, 2008), pp. 701–706.
29. D. Leyzberg, S. Spaulding, M. Toneva, B. Scassellati, The physical presence of a robot tutor increases cognitive learning gains. *Proc. Annu. Meeting Cogn. Sci. Soc.* **34**, 1882–1887 (2012).
30. J. Li, R. Kizilcec, J. Bailenson, W. Ju, Social robots and virtual agents as lecturers for video instruction. *Comput. Hum. Behav.* **55**, 1222–1230 (2016).
31. J. Kennedy, P. Baxter, T. Belpaeme, Comparing robot embodiments in a guided discovery learning interaction with children. *Int. J. Soc. Robot.* **7**, 293–308 (2015).
32. T.-H. Lin, S. Ng, S. Sebo, "Benefits of an interactive robot character in immersive puzzle games" in *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)* (IEEE, 2022), pp. 37–44.
33. D. Leyzberg, A. Ramachandran, B. Scassellati, The effect of personalization in longer-term robot tutoring. *ACM Trans. Hum.-Robot Interact.* **7**, 1–19 (2018).
34. B. F. Fuller, Y. Horii, D. A. Conner, Validity and reliability of nonverbal voice measures as indicators of stressor-provoked anxiety. *Res. Nurs. Health* **15**, 379–389 (1992).
35. F. Shaffer, J. P. Ginsberg, An overview of heart rate variability metrics and norms. *Front. Public Health* **5**, 258 (2017).
36. I. Pavlidis, N. L. Eberhardt, J. A. Levine, Seeing through the face of deception. *Nature* **415**, 35–35 (2002).
37. R. H. Good III, R. A. Kaminski, Assessment for instructional decisions: Toward a proactive/prevention model of decision-making for early literacy skills. *Sch. Psychol. Q.* **11**, 326–336 (1996).
38. L. A. Barnett, B. Storm, Play, pleasure, and pain: The reduction of anxiety through play. *Leis. Sci.* **4**, 161–175 (1981).
39. K. Charmaz, L. L. Belgrave, "Qualitative interviewing and grounded theory analysis" in *The SAGE Handbook of Interview Research: The Complexity of the Craft*, J. F. Gubrium, J. A. Holstein, A. B. Marvasti, K. D. McKinney, Eds. (Sage Publications, ed. 2, 2012), pp. 347–366.
40. R. Hoehn-Saric, D. R. McLeod, The peripheral sympathetic nervous system: Its role in normal and pathologic anxiety. *Psychiatr. Clin. North Am.* **11**, 375–386 (1988).
41. J. F. Thayer, F. Åhs, M. Fredrikson, J. J. Sollers III, T. D. Wager, A meta-analysis of heart rate variability and neuroimaging studies: Implications for heart rate variability as a marker of stress and health. *Neurosci. Biobehav. Rev.* **36**, 747–756 (2012).
42. A. Malliani, M. Pagani, F. Lombardi, S. Cerutti, Cardiovascular neural regulation explored in the frequency domain. *Circulation* **84**, 482–492 (1991).
43. F. Shaffer, R. McCraty, C. L. Zerr, A healthy heart is not a metronome: An integrative review of the heart's anatomy and heart rate variability. *Front. Psychol.* **5**, 1040 (2014).
44. H.-G. Kim, E.-J. Cheon, D.-S. Bai, Y. H. Lee, B.-H. Koo, Stress and heart rate variability: A meta-analysis and review of the literature. *Psychiatry Investig.* **15**, 235–245 (2018).
45. B. Vaishali, D. Vagish, V. Sricharan, S. Preejith, M. Sivaprakasam, "Statistical analysis of mental stress during oral presentation" in *2021 IEEE International Symposium on Medical Measurements and Applications (MeMeA)* (IEEE, 2021), pp. 1–6.
46. E. Farah, A. Desir, C. Marques, S. R. Hegde, A. Abreu, P. M. Polanco, C. Holcomb, D. J. Scott, G. Sankaranarayanan, Heart rate variability: An objective measure of mental stress in surgical simulation. *J. Surg. Educ.* **3**, 25 (2024).
47. J. Koenig, A. H. Kemp, T. P. Beauchaine, J. F. Thayer, M. Kaess, Depression and resting state heart rate variability in children and adolescents—A systematic review and meta-analysis. *Clin. Psychol. Rev.* **46**, 136–150 (2016).
48. N. Michels, I. Sioen, E. Clays, M. De Buyzere, W. Ahrens, I. Huybrechts, B. Vanaelst, S. De Henauw, Children's heart rate variability as stress indicator: Association with reported stress and cortisol. *Biol. Psychol.* **94**, 433–440 (2013).
49. S. Rotenberg, J. J. McGrath, Inter-relation between autonomic and HPA axis activity in children and adolescents. *Biol. Psychol.* **117**, 16–25 (2016).
50. J. Schmitz, B. Tuschen-Caffier, F. H. Wilhelm, J. Blechert, Taking a closer look: Autonomic dysregulation in socially anxious children. *Eur. Child Adolesc. Psychiatry* **22**, 631–640 (2013).
51. S. Ioannou, V. Gallese, A. Merla, Thermal infrared imaging in psychophysiology: Potentialities and limits. *Psychophysiology* **51**, 951–963 (2014).
52. B. Nhan, T. Chau, Infrared thermal imaging as a physiological access pathway: A study of the baseline characteristics of facial skin temperatures. *Physiol. Meas.* **30**, N23–N35 (2009).
53. C. A. O'Muircheartaigh, J. A. Krosnick, A. Helic, *Middle Alternatives, Acquiescence, and the Quality of Questionnaire Data* (Irving B. Harris Graduate School of Public Policy Studies, 2001).
54. I. Krumpal, Determinants of social desirability bias in sensitive surveys: A literature review. *Qual. Quant.* **47**, 2025–2047 (2013).
55. T. Holtgraves, Social desirability and self-reports: Testing models of socially desirable responding. *Pers. Soc. Psychol. Bull.* **30**, 161–172 (2004).
56. A. Bell, Designing and testing questionnaires for children. *J. Res. Nurs.* **12**, 461–469 (2007).
57. B. J. Peters, N. C. Overall, J. P. Jamieson, Physiological and cognitive consequences of suppressing and expressing emotion in dyadic interactions. *Int. J. Psychophysiol.* **94**, 100–107 (2014).
58. G. P. Kenny, R. McGinn, Restoration of thermoregulation after exercise. *J. Appl. Physiol.* **122**, 933–944 (2017).
59. I. Khalil, The effect of using repeated reading strategy on students' faulty pronunciation. *J. Al-Farahidi. Arts.* **13**, 440–448 (2021).
60. Misty Robotics, www.mistyrobotics.com/.
61. V. S. Vykhovanets, J. Du, S. A. Sakulin, An overview of phonetic encoding algorithms. *Autom. Remote Control.* **81**, 1896–1910 (2020).
62. V. Levenshtein, Binary codes capable of correcting deletions, insertions, and reversals. Sov. Phys. Dokl. 10, 707–710 (1966).
63. P. Boersma, D. Weemink, Praat: Doing phonetics by computer (2007); www.praat.org/.
64. Y. Jadoul, B. Thompson, B. de Boer, Introducing Parselmouth: A Python interface to Praat. *J. Phon.* **71**, 1–15 (2018).
65. M. E. Broome, A. P. Hellier, School-age childrens fears of medical experiences. *Issues Compr. Pediatr. Nurs.* **10**, 77–86 (1987).
66. N. Isnaini, T. Yunitasari, "Factors affecting anxiety level before examination electrocardiogram (ECG) in heart clinic" in *International Conference on Sustainable Innovation on Health Sciences and Nursing (ICOSI-HSN 2022)* (Atlantis Press, 2022), pp. 239–247.
67. M. Massin, G. Von Bernuth, Normal ranges of heart rate variability during infancy and childhood. *Ped. Cardiol.* **18**, 297–302 (1997).
68. C. L. Kaufman, D. R. Kaiser, J. Steinberger, D. R. Dengel, Relationships between heart rate variability, vascular function, and adiposity in children. *Clin. Auton. Res.* **17**, 165–171 (2007).
69. V. Shvartz, V. Danilov, S. Koroleva, E. Shvartz, S. Donakanyan, V. Permyakov, N. Erastova, E. Golukhova, "Become a Champion" Trial Investigators, Reference ranges of gender-and age-related heart rate variability parameters in Russian children. *Sci. Rep.* **15**, 5274 (2025).
70. D. Jelenova, M. Ociskova, J. Prasko, M. Hunkova, E. Karaskova, J. Kolarova, D. Vydra, V. Mihal, Quality of life in adolescents with inflammatory bowel disease and their parents--Comparison with healthy controls. *Neuroendocrinol. Lett.* **36**, 787–792 (2015).
71. Y. Wang, C. Zhang, S. Chen, X. Li, H. Jin, J. Du, Frequency domain indices of heart rate variability are useful for differentiating vasovagal syncope and postural tachycardia syndrome in children. *J. Pediatr.* **207**, 59–63 (2019).
72. P. Welch, The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Trans. Audio Electroacoust.* **15**, 70–73 (1967).
73. S. Tomar, Converting video formats with ffmpeg. *Linux J.* **2006**, 10 (2006).
74. G. Jocher, A. Chaurasia, A. Stoken, J. Borovec, Y. Kwon, K. Michael, J. Fang, Z. Yifu, C. Wong, D. Montes, Z. Wang, C. Fati, J. Nadar, Laughing, UnglvKitDe, V. Sonck, tkianai, yxNONG, P. Skalski, A. Hogan, D. Nair, M. Strobel, M. Jain, ultralytics/yolov5: v7. 0-yolov5 sota realtime instance segmentation, version 7.0, (Zenodo, 2022); https://zenodo.org/records/7347926.
75. P. Tsiamyrtzis, J. Dowdall, D. Shastri, I. T. Pavlidis, M. Frank, P. Ekman, Imaging facial physiology for the detection of deceit. *Int. J. Comput. Vis.* **71**, 197–214 (2007).
76. V. Braun, V. Clarke, *Thematic Analysis: A Practical Guide* (SAGE Publications, 2022).

Acknowledgments: We thank St. Thomas the Apostle School and the University of Chicago Laboratory Schools for collaboration and participation in this research. Vector art edited for use in Fig. 1 is attributed to S. Fahmi and zeimage508 on Freepik. **Funding:** This work was supported by the National Science Foundation awards 2312352, 2312353, 2312354, 2202802, 2202803, and 2247381. **Author contributions:** L.L.W., J.E.M., B.M., and S.S. were responsible for conceptualizing the study and identifying research questions. The experiment design was led by L.L.W., with input from J.E.M., B.M., and S.S. The robot implementation was done by L.L.W. The tablet application was designed by P.V. Running the study was led by L.L.W. with assistance from P.V. Data analysis was led by L.L.W. with input from S.S. and the assistance of P.V. and a team of undergraduate research assistants. Writing the paper was led by L.L.W. with revisions and contributions from S.S., J.E.M., and B.M. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** Deidentified data with code used to evaluate the conclusions and code used to run the study can be found at https://doi.org/10.5281/zenodo.16739387.

Submitted 15 November 2024
Accepted 19 August 2025
Published 10 September 2025
10.1126/scirobotics.adu5771