ABSTRACT
This paper studies how eye-tracking can be used to measure and facilitate joint attention in parent-child interaction. Joint attention is critical for social learning activities such as parent-child shared storybook reading. There is a disassociation of attention when the adult reads texts while the child looks at pictures. We hypothesize the lack of joint attention limits children’s opportunity to learn print-related skills. Traditional research paradigm does not measure joint attention in real-time during shared storybook reading. In the current study, we simultaneously tracked eye movements of a parent and his/her child with two eye-trackers. We also provided real-time feedback to the parent where the child was looking at, and vice versa. Changes of dyads’ reading behaviors before and after the joint attention intervention were measured from both eye movements and video records. Baseline data show little joint attention in parent-child shared book reading. The real-time eye-gaze feedback significantly changes parent-child interaction and improves learning.

Author Keywords
Joint attention, eye-tracking, eye gaze, inter-person interactions, shared storybook reading.

ACM Classification Keywords

INTRODUCTION
Joint attention is critical in social learning such as parent-child shared storybook reading. To acquire print-related skills, children and adults must coordinate their joint attention during dynamic interactions [6, 11, 14, 22]. This requires the partners to maintain a triangular attentional structure, i.e., attending simultaneously to the target of learning and among themselves. However, prior eye-tracking studies show that pre-reading children focus almost exclusively on illustrations while parents read from print texts [4, 5, 8, 9]. In fact, it is difficult for a child to know where on the page the adult reader is attending to, and vice versa. This limits the partners’ ability to regulate their joint attention. We hypothesize that the lack of joint attention impedes the acquisition of print-related skills, and children will learn better if we can facilitate the regulation of joint attention.

The present study has two goals. First, we seek to objectively measure the joint attention in shared storybook reading, by simultaneously tracking the eye gaze of the parent and the child. Second, we investigate whether two eye-gaze-based interventions enhance parent-child joint attention during reading. The interventions target the fact that partners in shared reading do not know where the other person is attending to. One intervention involves showing a real-time moving cursor on the child’s monitor that indicates where the parent is looking. The other intervention shows the parent where the child is looking. With this critical piece of information, it is hypothesized that the dyad can better regulate their joint attention, which will facilitate children’s learning of new sight words.

RELATED WORK
Print-related skills refer to children’s knowledge of the rules for translating the particular writing they are trying to read into sounds, including letter knowledge, phonological awareness, knowledge of letter-sound correspondence, print knowledge, and sight word recognition [18, 20, 24]. There is converging evidence that a key to develop print-related skills is to engage children in a joint attention on print words. This can be achieved by pointing to words while reading and having print-focused conversations [8, 10], both of which are key elements in a joint attentional
interaction. However, traditional reading strategies have some inherent limitations. One problem is that the two partners do not have accurate knowledge of where the other person is attending to at any moment, which makes attention regulation difficult and ineffective. In addition, most existing reading strategies are adult-centered - i.e., the parent regulates the child’s attention without much information about the child’s interest at the moment.

We conjecture the ideal solution to the lack of joint attention in the naturalistic shared reading is to provide partners consistent, individualized, and real-time feedbacks during reading. By externalizing adults’ reading processes, the pre-reading children will have a model which they can mimic and internalize. And with a projection of children’s attention and thinking process, adults will be able to scaffold and strategize pedagogical goals accordingly.

Short of a magic window into each other’s mind, the state of eye-tracking technologies allows us to show where the partner is looking at in real-time. We can show parents or children, a cursor on the computer screen that corresponds to the gaze location of the other person. This process of helping people understand real-time eye movements is called the eye-gaze awareness training. Eye movements provide critical information that is missing in the traditional shared reading task.

First, the location of the eye gaze indicates the focus of attention at any given moment [15, 16]. As previous studies showed that the coordination of joint attention is essential for communication, collaborative visual search and problem solving tasks [1, 12, 17], we expect that discovering children’s real-time attention state may also trigger adults’ regulations of joint attention during shared book reading. The real-time eye gaze information is more instructional to the pre-reading children, who will see where and how grown-ups look when they read. We expect that children will be more likely to repair their own flawed mental models if they recognize that theirs conflict with adults’ reading model. Children may even start to follow adults’ scanning patterns and pay more attention to texts, which in turn provides more teachable moments for parents to introduce print-related knowledge and skills.

Second, the scanning pattern of the eyes can be used to gauge children’s cognitive processes. For example, recent psycholinguistic studies have shown that the eye typically follows the order of words in sentence comprehension [2, 7, 21]. It is intuitive that a child is not paying attention or is having difficulty comprehending a story if his or her eyes do not keep up with things mentioned by the adults. Adults may slow down the pace of reading and help children catch up in this situation.

Lastly, having access to the other partner’s eye movements may change the dynamics of shared reading as well as greatly reduce the time and energy that two partners spend on reengaging joint attention. Gone is the need to ask “Are you looking at here?” because the answer is on the screen.

The success or failure of a pedagogical attempt is immediately seen on the screen as well. Adults can give children more prompt and precise feedback when they watch children’s real-time eye movements. If a parent has the goal of teaching print but notices that his/her child’s eye movements only focus on pictures, the parent probably would not ignore this information and keep talking about the print by himself/herself. Instead, the parent may first follow the child’s interest on pictures and then utilize verbal or nonverbal strategies to direct the child’s attention to texts.

**METHODOLOGICAL INNOVATIONS**

The current study involves a series of methodological innovations. While there are a handful of published studies looking at children’s eye gaze during shared book reading, none has investigated the correlation and contingency between eye movements of children and parents. Using two eye trackers simultaneously, we tracked dyads’ eye movements and measured their joint attention in real-time during shared book reading. The data and methodology will be useful to a wide array of researchers interested in joint attention and collaborative behaviors.

Furthermore, we sought to intervene in the dynamic interactions by showing dyads the other partner’s eye gaze positions. To improve pre-readers’ understanding of reading, we presented a moving cursor on the computer screen which shows where adults are looking at any moment. We also investigated how parents use the real-time eye gaze information to regulate joint attention.

**EXPERIMENT**

There are three experiments in the current study. Experiment 1 is the baseline in which we tracked the eye gaze of both the parent and the child during the naturalistic shared reading. We define joint attention as when the partners look simultaneously at (or near) the same visual object on a page (see the examples in Figure 1). Children’s sight word recognition was measured before and after the reading. We anticipate little joint attention on print and
therefore limited print learning. This serves as the control condition for subsequent intervention experiments.

Experiment 2 and 3 are two intervention experiments in which we investigated whether the real-time feedback of eye movements enhances parent-child joint attention and children’s print-related learning. We presented children how their parents read texts in Experiment 2 and showed parents their children’s real-time eye movements in Experiment 3. We hypothesize that the new paradigm will help dyads regulate joint attention and help children learn reading. The research methods of the three experiments are illustrated as follows.

**Participants**
Thirty-seven dyads participated in Experiment 1; they also serve as the comparison group for the subsequent intervention experiments. Experiment 2 involved twenty-seven parent-child dyads. Experiment 3 involved twenty-eight dyads. All children participants were 4-5 year old English speakers who had no history of hearing, vision, or cognitive impairments. Parent participants were required to be person who reads most frequently with children at home.

**Materials**
Three age appropriate storybooks were presented for dyads to read in all three experiments. Children’s sight word learning was measured before and after reading by asking children to name content words sampled from the storybooks. We used stuffed animals as props to ask children where adults look on a page (pictures vs. texts) and in which direction they read (left to right vs. right to left).

**Apparatus**
Two contact-free eye trackers, a Tobii X50 (see http://www.tobii.se) and an Eyelink 1000 system (www.eyelinkinfo.com), were used in the study. The Tobii X50 system is a video-based remote eye tracking system that makes no contact with the participant. It samples at 50Hz, and has a typical accuracy of approximately 1 visual degree (measured by repeated calibrations). The system uses infrared cameras to automatically capture eye images from a reading distance. Eyelink 1000 is also an infrared-based system, but with much higher accuracy (0.5 visual degree) and sampling rate (500 Hz). As a remote system, it also allows contact-less operations. For each dyad, the parent was eye tracked by Tobii X50 and the child was eye tracked by Eyelink 1000. Two video recorders were used to record verbal and non-verbal interactions among the dyad.

For each dyad, the parent and the child sat across a child-sized table at a 90 degree angle. One LCD monitor (1280x1024 pixels resolution) and Eyelink 1000 were put approximately 60cm away from the child; while another LCD monitor (1280x1024 pixels resolution) and Tobii X50 were put approximately 60cm away from the parent (see Figure 2 for details of the set-up). Stimuli were presented simultaneously on both monitors. Stimulus presentation and eye movement calibration and recording were done using the Double Tracker program developed in our lab. The data were then exported offline for statistical analyses.

The child and the parent were individually calibrated on the Eyelink and Tobii eye trackers. They were monitored throughout the study and recalibrated as necessary. Both remote eye trackers use the corneal reflection to compensate for head movements. Our past experience shows that neither exhibits substantial drifts that affect results in our experimental paradigms.

To the extent both eye trackers are accurately calibrated, aligning the gaze positions is straightforward. Because identical images are shown on both monitors, we took the gaze position in terms of screen coordinates and mapped to the other screen. This is essentially the same as the standard technique of displaying a gaze cursor on the same monitor as the participant is viewing.

**Procedures**
For all three experiments, each parent-child dyad read 3 books in four reading trials (order counter-balanced among participants). They read one same storybook in the first and fourth trial, and the other two storybooks in the second and third trial. In the fourth trial the parent was asked to teach three words that the child did not recognize based on the pretest.

In Experiment 1 parents read children storybooks on screens but there was no eye gaze feedback for both partners. Experiments 2 and 3 involved showing a moving cursor on one of the participants’ screen; the moving cursor indicated the location of the other person’s eye gaze in real-time. We ensured children and parents understood the gaze indicator using an iSpy-like game. Even the youngest children had no problem understanding the correspondence. In Experiment 2, we showed the child where the parent was looking. In Experiment 3 the parent saw where the child was looking. In both conditions the other participant looked at a normal, static display of the page. To tease apart the
impact of the instructions from the moving cursor, we asked children in all three experiments to follow the parent’s eye gazes while listening to stories, even though they could not actually see the eye gazes in Experiment 1 and 3.

Data Transcription and Coding
For the eye movement data, areas of interest (AOIs) were defined for pictures and texts on each page of the three books. An AOI includes a margin of approximately 1.3 cm on each side of object it encloses. Eye movements fell outside of any AOIs (e.g., on the white background) were excluded from analyses. Data were also discarded when the eye trackers lost track, which could occur when the participant looked away from the monitor, moved the head rapidly, or closed his/her eyes. The average percentage of children’s fixations on text AOIs in the first reading trial was compared with that in the fourth reading trial.

To measure the real-time joint attention during the reading session, we compared the distance of two partners’ eye gaze locations with a cut-off value of 201.18 pixels which was determined for three reasons. First, 201.18 pixels correspond to the 80th percentile of the effective eye movement data points in the distance distribution of joint attention trials in the pretest. Second, the visual angle which corresponds to the 201.18 pixels is about 10 degrees (Eyelink systems typically have 20 pixels / degree). The human fovea, where we have clear vision, is about 2 degrees. So the visual angle of 10 degrees is not a too small window size for a definition of joint attention. Third, the 201.18 pixels are close to the size of two 5-letter-long print words in pixels (the average length of a 5-letter-long word is 100 pixels). Therefore, we believe this is a very reasonable window to define joint attention in reading.

We determined the joint attention exists if the distance is smaller than 201.18 pixels and does not exist if the distance is larger than or equal to 201.18 pixels. The percentage of time when the distance of two partners’ eye gaze locations is smaller than 201.18 pixels represents how much joint attention the dyad has when reading together.

Video recordings of the parent-child shared book reading interactions were transcribed and coded with the InqScribe software. Adapting from the coding systems in previous studies [3, 13, 19, 23], we have developed a coding system to analyze 11 types of parent-child joint attention interactions, including parents looking at children, children looking at parents, parents’ verbal attention regulation, parents asking children to look at specific words, parents teaching specific words, parents providing children specific feedback, parents pointing to words on the screen, children reading texts along with parents, children pointing to texts on the screen, children completing a sentence with parental prompt, and children talking about print.

RESULTS AND DISCUSSION
We hypothesized that (a) there is limited parent-child joint attention to texts in the naturalistic shared storybook reading, and (b) the eye-gaze feedback facilitates joint attentional regulation and improves the acquisition of print skills. Our data support both predictions.

Specifically, Experiment 1 showed that the average percentage of time children had joint attention with parents on texts in reading trial 1 was very small (2.91%). In the reading trial 4, when adults were asked to teach children three keywords, children significantly increased their joint attention on texts to 6.41% ($t(36) = 2.48, p = .018$). Children learned an average of 0.38 words as measured by pre- and post-test of word recognition.

Both interventions significantly increased parent-child joint attention on texts. In Experiment 2, where children saw parents’ eye gaze, the joint attention increased from 5.35% in reading trial 1 (no-intervention trial) to 22.7% in reading trial 4 (intervention trial). The increase of 17.35% is significantly higher ($p = .000$) than the increase of 3.5% in Experiment 1 (see Figure 3).

![Figure 3: Percentages of parent-child joint attention on print from reading trial 1 to 4 between Experiment 1 and 2](image)

More print-directed joint attention resulted more word learning: children learned an average 1.0 word as compared to 0.38 words in Experiment 1, $p = .02$. This result indicates that children who received the eye-gaze awareness training learned more keywords from the pretest to the posttest than children who did not receive this training.

Moreover, seeing parents’ eye movements also induced conceptual changes. Children were at chance in the pretest answering questions on where adults look in reading (62.96%) and which way they scan the text (51.85%). After the intervention the correct rate increased to 88.89% and 81.48%, respectively, and both are significantly above chance ($p = .000$ and $p = .001$ respectively).

Positive effects are also found in Experiment 3, in which the parent received real-time feedback on the child’s visual attention but the child did not see the parent’s eye gaze. Joint attention on texts increased from 3.48% in reading trial 1 (no-intervention trial) to 12.87% in reading trial 4 (intervention trial). The increase of 9.39% was significantly...
higher than the increase of 3.5% in Experiment 1 \((p=.012, \text{ see Figure 4})\).

Parents became more effective facilitating children’s sight word learning. Children learned 1.25 words, significantly higher than that in Experiment 1 (0.38 words, \(p=.000\)). This result further confirmed that children learned more sight words from the pretest to the posttest when parents used the eye-gaze feedback to effectively direct children’s attention to words.

For the behavior analysis, the frequencies (times per minute) of these 11 parent-child interaction behaviors in reading trial 4 for all the three experiments are shown in Table 1.

<table>
<thead>
<tr>
<th>Types of Behaviors</th>
<th>Experiment 1 (baseline)</th>
<th>Experiment 2</th>
<th>Experiment 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parents look at children</td>
<td>1.46(2.35)</td>
<td>.69(1.22)</td>
<td>1.56(1.82)</td>
</tr>
<tr>
<td>Children look at parents</td>
<td>2.48(2.86)</td>
<td>1.33(2.18)</td>
<td>1.97(2.24)</td>
</tr>
<tr>
<td>Parents’ verbal attention</td>
<td>.26(.38)</td>
<td>.17(.34)</td>
<td>.78(1.1) *</td>
</tr>
<tr>
<td>Parents ask children to look at specific words</td>
<td>.53(.76)</td>
<td>.24(.54)</td>
<td>1.56(1.32) **</td>
</tr>
<tr>
<td>Parents point to words on the screen</td>
<td>.15(.53)</td>
<td>.01(.06)</td>
<td>.22(.61)</td>
</tr>
<tr>
<td>Parents teach specific words</td>
<td>.17(.44)</td>
<td>.08(.18)</td>
<td>.74(.88) **</td>
</tr>
</tbody>
</table>

| Parents provide specific feedback      | .01(.04)                | 0(0)         | .62(.72) **  |
| Children read along with parents       | .25(.42)                | .61(.93) *   | .27(.55)     |
| Children complete a sentence with parents’ prompt | .32(1.15)              | 1.14(2.21) * | .27(.67)     |
| Children point to texts on the screen  | .13(.36)                | .13(.33)     | .39(.61)     |
| Children talk about print              | .27(.5)                 | .37(.67)     | 1.26(1.1) ** |

* \(p<.05\), ** \(p<.01\)

Table 1: Mean (standard deviation) of the behavior occurrences per minute in reading trial 4 for all 3 experiments

To compare the changes of the frequencies of the behaviors from reading trial 1 to reading trial 4 between Experiment 1 (control group) and Experiment 2 (child eye-gaze awareness training group), we did repeated measures ANOVAs using the average frequency of each behavior as the within-subjects variable (reading trial 1 vs. reading trial 4) and whether children received the eye-gaze awareness training as the between-subjects factor (child eye-gaze awareness training group vs. control group).

The results suggested that children in both groups significantly increased the occurrences of the behavior of reading texts along with parents from the first to fourth reading trial, but children who received the eye-gaze awareness training showed a significantly larger increase \((p=.041)\). For children’s completing a sentence with parental prompt type of behavior (e.g., children finished reading the last word of a sentence after parents stopped reading and waited for children’s responses), children in both groups increased the occurrences of this behavior from the first to fourth reading trial, but children who received the eye-gaze awareness training showed a significantly larger increase \((p=.046)\).

Overall, the comparisons between Experiment 1 and 2 indicated that with the eye gaze direction more tightly tied to the focus of joint attention, children saw an external representation of reading processes unfolding in real-time and therefore responded more to parents’ word teaching attempts, such as reading texts along with parents or reading the last word of a sentence with parental prompt. Children’s increased responses to parents’ reading
strategies further improved children’s visual attention to print as well as their print learning outcomes.

We did the similar comparisons for the changes of the frequencies of the behaviors from reading trial 1 to reading trial 4 between Experiment 1 (control group) and Experiment 3 (parent eye-gaze awareness training group). For parents’ verbal attention regulation type of behavior (e.g., adults saying “Look at the screen/words.”), parents in both groups increased the occurrences of the behavior of verbally asking children to attend to print from the first to fourth reading trial, but parents who received the eye-gaze awareness training showed a significantly larger increase ($p=.000$). For the parents asking children to look at specific words type of behavior (e.g., “Can you help me find the word ‘cat’ on the screen?”), parents on average increased the occurrences of this behavior from the first to fourth reading trial, but parents who received the eye-gaze awareness training showed a significantly larger increase ($p<.001$). For parents’ teaching specific words type of behavior (e.g., adults saying “The word ‘book’ starts with a letter ‘B’ and ends with a letter ‘K’; “The first word on the second line starts with a ‘buh’ sound, can you try to sound it out?”), parents on average increased the occurrences of this behavior from the first to fourth reading trial, but parents who received the eye-gaze awareness training showed a significantly larger increase ($p=.002$). For parents’ providing specific feedback type of behavior (e.g., adults saying “Yes, you are looking at the right place.”; “No, you are not looking at the place I want you to look.”), parents in the eye-gaze awareness training group significantly increased the occurrences of the behavior of providing children the specific feedback from the first to fourth reading trial ($p=.000$), but parents in the control group did not change much of this behavior.

For children’s talking about print-related things type of behavior, children in both groups increased the occurrences of the behavior of asking or answering print-related questions from the first to fourth reading trial, but children whose parents received the eye-gaze awareness training showed a significantly larger increase ($p=.000$).

The above comparisons between Experiment 1 and Experiment 3 suggested when parents received the real-time feedback of their child’s visual attention during shared book reading, they were better at regulating children’s attention and at providing specific feedback following children’s responses. Many parents spontaneously adjusted their reading strategies in response to children’s real-time attention states. There were a lot more attention regulation types of interactions in the parent eye-gaze awareness training group, such that parents verbally asked children to pay attention to texts and constantly pointed to some specific words on the screen. Since parents could see where children looked at after a request or question, the redundancy of frequently asking children “Are you following mom and looking at the place I talked about?” was significantly reduced. Instead, parents directly provided children with more prompt and precise feedback such as “That’s great, I can see you are looking at the right word.” These reading behaviors are more efficient for accelerating communications between parents and children.

In addition, parents in the eye-gaze awareness training group showed more frequent and effective print teaching behaviors compared to parents in the control group, such as asking questions about particular letters or sound within a word, helping children spell or sound out a word, and commenting on print-related contents. Parents in the control group also used some strategies to teach print, but since they did not know where children were looking at, their reading strategies were neither effective nor efficient. For example, the control group parents usually looked at their children and talked about one specific word for a long period of time while their children already got bored and did not really look at that particular word. As a consequence children could hardly recognize that same word in the posttest. In contrast, with the help of real-time visual attention feedback, parents in the eye-gaze awareness training group had better opportunities to observe their child’s attention state and fine-tune their interactions to increase child interest and participation. Specifically, these parents could reduce the time and energy spent on looking at children or verbally checking children’s responses. Instead, they frequently and naturally incorporated varies efficient and well-organized literacy activities within children’s short attention span. And children more readily learned those words when they had more frequent but shorter durations of learning experiences.

Similarly, children who experienced positive direction, coaching, and correction more easily attended to and internalized the information and skills that parents attempted to teach them, and developed the interest and motivation to sustain their learning. These changes in turn provided more teachable moments for parents.

**CONCLUSION**

The current study measured parent-child joint visual attention in real-time, which allows us to go beyond prior research that focuses exclusively on the child in shared book reading, and study shared reading as a joint attentional interaction involving dynamic transactions between partners and real-time cognitive strategies within individuals. The data and methodology of this study would also be useful to a wide array of research topics on collaborative learning and interaction situations.

The new technology enabled us to investigate a number of issues in shared storybook reading. Building on prior research, we found that pre-reading children have limited joint attention with their parents while listening to storybooks. More importantly, parents typically have little
information about where children are attending to, and children have even less idea about how adults actually read. This results in a poorly regulated joint attentional interaction when it comes to learning print-related skills.

An important contribution of the present study is the intervention experiments, in which we remedy the joint attentional structure by leveraging the eye-tracking technology. By providing real-time feedback of the partner’s visual attention, we demonstrate significant improvements in the amount of joint attention on print texts and changes in parental attentional regulation strategies during reading. More interestingly, children did not simply look at the moving cursor or print words, but actually read and processed the words. This was shown by increased word learning by children, along with children’s changes of concept of reading processes. Our intervention targets limitations in joint attention regulation in the traditional concept of reading processes. Our intervention targets the extent learning involves joint attention (e.g., in math tutoring), the eye-gaze feedback may be an effective aid for learning. More importantly, data suggest that by providing a critical piece of information -- namely, where the partner is looking -- we can facilitate the regulation of joint attention and improve children’s acquisition of print-related skills.

REFERENCES
