A Comparative Study of Visualizations with Different Granularities of Behavior for Communicating about Autism

HA-KYUNG KONG*, JOHN LEE*, and KARRIE KARAHALIOS,
University of Illinois at Urbana-Champaign

This paper presents a comparative study of two webtools developed to capture engagement via visualization of coordinated communication behavior in children with autism. A clear preference arose for different tasks based on behavior granularity emphasis in the two visualizations. The survey and interview results further revealed the importance of showing behavior patterns, rather than displaying a single behavior without context, in behavior visualization. Based on the results, we propose three granularity-related features to incorporate into behavioral visualizations for communication in clinical settings: separating modalities, coordinating dyadic interactions, and displaying micro-behaviors.

CCS Concepts: • **Human-centered computing** → **Visualization design and evaluation methods**; • **Applied computing** → *Health care information systems*;

Additional Key Words and Phrases: Dyadic interaction; behavioral visualization; joint attention; behavioral imaging; autism; webtool design

ACM Reference format:
Ha-Kyung Kong, John Lee, and Karrie Karahalios. 2017. A Comparative Study of Visualizations with Different Granularities of Behavior for Communicating about Autism. *Proc. ACM Hum.-Comput. Interact.* 1, 2, Article 59 (November 2017), 16 pages.
https://doi.org/10.1145/3134694

1 INTRODUCTION

Imagine a scenario where a clinician pre-screens a child for autism; immediately after the pre-screening session, a visual representation of the session becomes available to the clinician. Previous works have explored this idea, and showed that visualizations of clinician-patient communicative behavior are effective in uncovering outlier behavior patterns [7, 15, 17]. People additionally suggested that visualizations would be useful for facilitating communication between individuals by providing a common resource for communication and collaborative interaction.

Communication with parents is an essential part of Autism Spectrum Disorder (ASD) diagnosis and treatment. Osborn and Reed found that parents of children diagnosed with ASD anticipated more debriefing of the diagnosis session and general education about ASD after a child-clinician session, yet did not receive it and became frustrated [22]. They specifically expressed interest in the

*Both authors contributed equally to the paper

This work is supported by the National Science Foundation, under grant CCF-1029679.
Authors' addresses: H. Kong, J. Lee, K. Karahalios Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL 61801, US; .

2573-0142/2017/11-ART59 $15.00
https://doi.org/10.1145/3134694

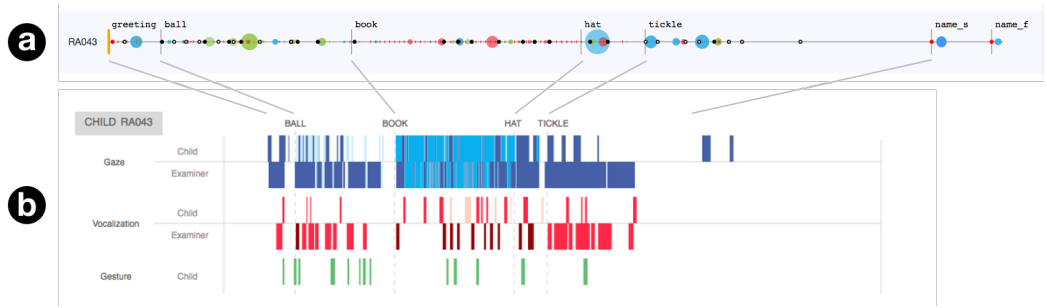

Fig. 1. (a) Plexlines and (b) an EnGaze visualization of the same child in a RABC session. The grey lines connecting the two visualizations indicate the start of different phases in a single session. While the Plexlines visualization accentuates behavior with longer duration, the EnGaze visualization is capable of showing more micro-behaviors. We describe the details of the Plexlines and EnGaze visualizations in Figures 3 and 7.

idea of "broad information sheets [provided] to parents at the time of diagnosis" [22]. Clinicians can leverage visualizations, such as the ones shown in Figure 1, to fulfill their wants by stepping through the reasoning of the diagnosis with rich contextual information.

Although previous work showed that visualizations could be used for communication in clinical settings, their main focus has been on using visualizations to find reoccurring patterns and detect outliers [15, 17]. Our work extends this literature to uncover features of visualizations that are useful for communicating about ASD through a comparative study of two behavior visualizations. There are advantages in conducting a qualitative study with multiple visualizations. The first strength of a comparative study is that people are less reluctant to criticize when there are multiple designs, and provide less inflated ratings [26]. Second, the contribution of each feature is hard to assess in a single visualization study since a combination of features makes a visualization useful. By comparing two visualizations intended for the same purpose, we can have a better sense of which features are useful for a certain task by asking the participants to choose a single visualization for a use case and prompting for specific features. Thus, we conducted a comparative study on two visualizations, Plexlines and EnGaze, which summarize an autism pre-screening session in a single visualization. We discuss the details of the visualizations and the reason for choosing these two visualizations in the following section.

Through the interviews, we found that the tasks clinicians wanted to perform with behavioral visualizations fell into two categories: 1) single session tasks and 2) multi-session tasks that involved comparison of multiple sessions. Participants expressed a clear visualization preference for each category – EnGaze for single session tasks and Plexlines for multi-session tasks. This preference arose from different granularities emphasized in the visualizations. Plexlines focused on low granularity behaviors through emphasizing behaviors of longer durations; EnGaze involved high granularity behaviors through showing behaviors of shorter lengths. We present the design decisions of each visualization that led to different granularities in the results section. Based on the interview results, we discuss three most commended features that help communicate the results of a clinical session: separating modalities (e.g., vocal, gesture, gaze), coordinating dyadic interactions (opposed to visualizing only the child), and displaying micro-behaviors. We further envision a broader use of behavior visualization for communication in other clinical settings and explore potential use cases in the discussion section.

2 RAPID ABC

The two visualizations used in our comparative study use data from an autism pre-screening child-examiner play protocol called the Rapid ABC (RABC) [23]. The RABC was developed to identify infants and toddlers at-risk for ASD. An RABC session consists of 5 stages of play: greeting, rolling a ball, reading a book, wearing the book as a hat, and tickling. The stages are highly structured for easy comparison across all participants. In each stage, the examiner initiates interactions with the child to elicit a response. For example, the examiner might ask the child: "Look at my book!" or "Where is the duck?" during the book stage. Upon completion of an RABC session, the examiner then evaluates the child's responses based on the level of engagement the child displayed. The dataset provides detailed hand-coded annotations of the child and examiner's communicative behaviors in three modalities—gaze, gesture, and vocalization—along with the corresponding video. Currently the videos are hand-coded, but colleagues at our and other research institutions are working on automating this process through computer vision techniques.

3 PLEXLINES AND ENGAZE

We selected Plexlines and EnGaze [15, 17] for our comparative study for two specific reasons. (1) Plexlines and EnGaze were based on the same dataset and visualized the same set of behaviors, which made it easier to make a fair comparison. (2) The two visualizations were both designed to assess engagement at a glance, but they used two different design approaches to achieve this goal. TipoVis is another visualization that makes use of the same dataset [7]. It excelled at identifying correlations or the lack thereof between two specific behaviors, such as pointing and vocalizing of a child, but it lacked the ability to visualize all modalities of behavior unlike Plexlines and EnGaze. Since TipoVis is limited to two behaviors, any co-occurring behavior patterns that involve three or

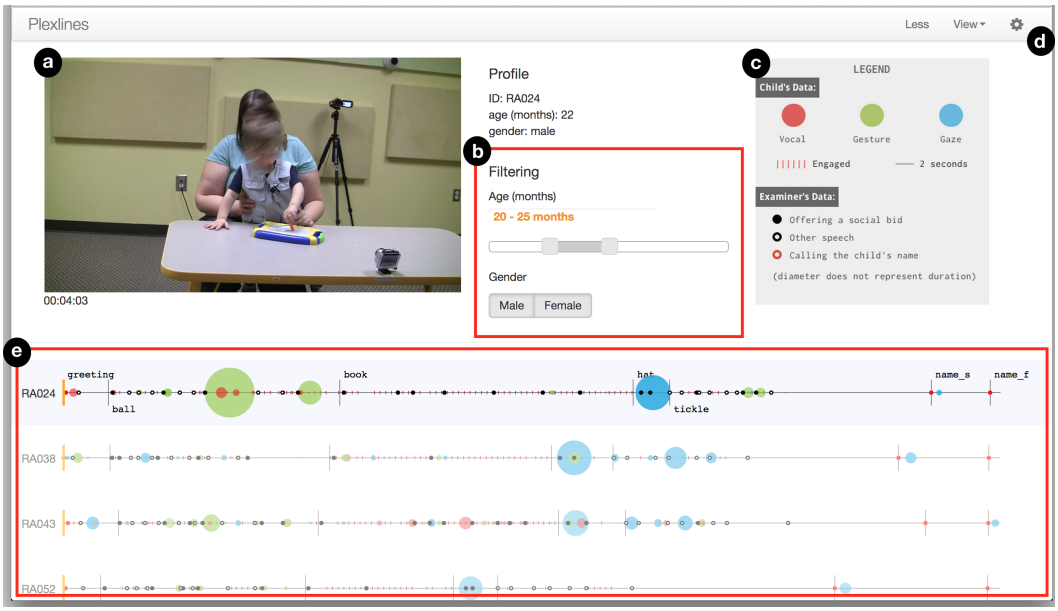

Fig. 2. The main view of the Plexlines visualization tool. Through this interface, users can select a visualization from (e) and view the recorded video corresponding to the visualization in (a). The visualization tool provides (b) filtering options, (c) the legend, and (d) customization features (once the user clicks the cog).

more behaviors cannot be displayed in TipoVis. Thus, we chose Plexlines and EnGaze since they provide a more comprehensive overview of a session at a glance.

Plexlines uses color coded circles on a timeline to visualize three different modalities of child behaviors (See Figures 2 and 3). Plexlines was designed to display communicative behaviors and to aid in identifying developmental delay in the domain of ASD. The diameter of the circle is proportional to the duration of the annotated behavior, which accentuates behaviors with longer duration, such as prolonged eye contact with the examiner. The visualization also offers additional features for its users, such as providing an aggregate view to see patterns over multiple visualizations (see Figure 4b) and allowing users to customize the visualization by removing certain behaviors based on their specific needs. The details of the customization capabilities such as filtering, sorting, and hiding certain behaviors can be found in [17]. A user study conducted with ASD researchers and clinicians revealed that Plexlines has potential to be integrated into existing behavioral evaluation processes, aid in the detection of developmental delays in young children, and serve as a visual artifact to better communicate with parents.

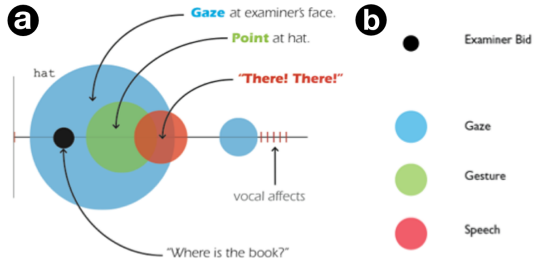

Fig. 3. (a) A short child-examiner interaction depicted in Plexlines with the (b) legend. In (a), the examiner asked, "where is the book?" and the child response to the examiner by pointing at the book (green circle) and saying "there, there" (red circle), while maintaining eye contact (blue circle)

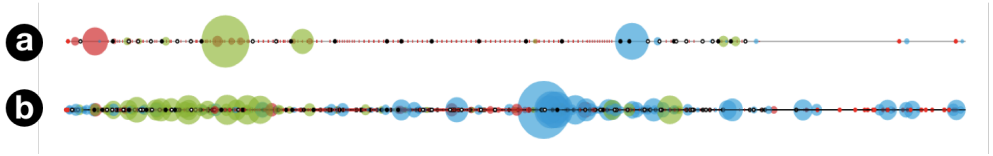

Fig. 4. (a) One Plexline compared against an (b) aggregate of the same sequence of events.

Similarly, EnGaze visualizes communicative behaviors, but by using colored rectangles on a timeline (See Figures 5 and 6). The unique characteristics of EnGaze is the inclusion of detailed examiner's behaviors and its focus on identifying moments of joint attention during an RABC session. Joint attention is a set of communicative behaviors that signals the shared focus of two individuals, and the lack of joint attention is a defining characteristic of ASD. Thus, identifying joint attention, or the lack thereof, can lead to early detection of ASD, and in turn contribute to early intervention. EnGaze provides several features that highlight moments of joint attention. Users can highlight moments when the child responds to the examiner (See Figure 7), and customize the visualization through hiding user selected behaviors in the visualization. Unlike Plexlines, EnGaze did not emphasize behaviors with longer duration. Rather, it made it easier for users to observe behaviors with shorter duration, such as exchange of glances, which often occurs in a fraction

of a second. In our work, we define such short behaviors as micro-behaviors. Many important communicative features are micro-behaviors, but micro-behaviors are easily overlooked in everyday communications.

Despite the differences between EnGaze and Plexlines, the two works received similar feedback. The main findings for EnGaze overlapped with that for Plexlines — it can aid clinicians in identifying children with developmental delays. There were four proposed use cases for the two visualization tools overall.

- Identifying children with developmental delays
- Providing an objective, quantifiable baseline for their experiences during an assessment
- Showing and defining characteristics of ASD for education
- Communicating with parents by using the visualization to provide feedback to parents.

In this paper, we delve deeper into last two use cases that relate to communication in clinical settings.

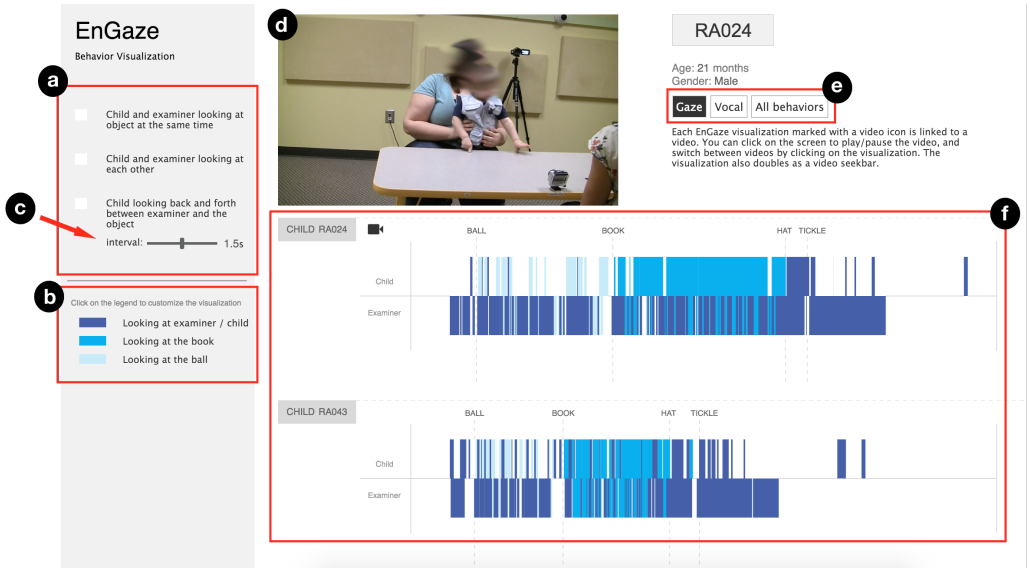

Fig. 5. The main view of the EnGaze visualization tool. This view is currently only showing the gaze of the child and examiner as selected in (e). Similar to the Plexlines visualization tool, users can view the video through (d) corresponding to a visualization in (f). Users can highlight moments of attention through the options in (a). The legend is provided in (b).

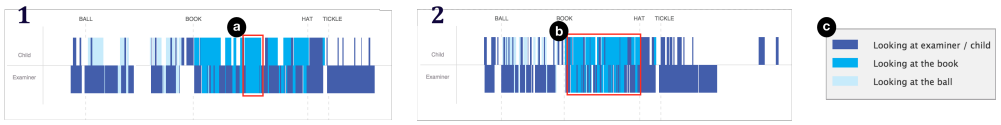

Fig. 6. Two EnGaze visualizations, showing only gaze behaviors. Each EnGaze visualization consists of a child section (top) and a examiner section (bottom). This figure currently highlights two behaviors discovered through the visualization. In the first EnGaze visualization (a), both the child and examiner are looking at the book. In the second visualization (b), the examiner and child are displaying gaze alternation, which is a strong indicator of joint attention. The legend is provided in (c).

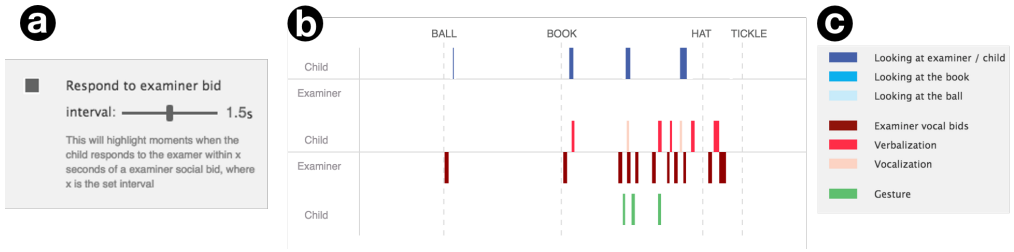

Fig. 7. EnGaze visualization showing all three modalities of behaviors, and highlighting only the child's responses. This can be accomplished by enabling the button shown in (a), "Respond to examiner bid" and selecting a time window for a child responses. (b) The EnGaze visualization is currently showing all three behaviors, with gaze on top, vocalization in the middle, and gesture on the bottom. Only the child's actions exist for gestures. Following the legend (c), examiner prompts are marked as dark red in the middle. All other colored bars are direct responses from child.

4 RELATED WORK

Our work on studying behavioral visualizations for effective communication builds on past research on visualizations of behavior and visualization for communication that we discuss in this section.

4.1 Visualizations of Behavior

Previous research in visualization of human behavior explored visualizing different modalities of human behavior in a specific domain of interest. Researchers have explored visualizing conversation (e.g. [1, 6]), eye gaze (e.g. [2]), and multimodal sensory input data (e.g. [13, 14]) in both clinical and non-clinical contexts. More recent work by Lee et al. and Kong et al. showed that customizable visualization tools for structured child-examiner sessions can aid clinicians in detecting children with developmental delay [15, 17]. Han et al. used the same dataset to create TipoVis, a behavior analysis tool, which allows users to select and visually explore the temporal patterns of two specific social and communicative behavior. In two separate case studies, the visualization enabled researchers to identify correlation between two behaviors [7]. Each of these behavioral visualization tools were created with a specific purpose, yet several previous works have uncovered additional use cases of visualizations beyond their intended application. Kim et al. anticipated her data analysis and visualization tool, BEDA, to facilitate finding patterns across various sensory inputs for children at risk, and unexpectedly found that BEDA also increased behavioral scientists' confidence in conducting analyses of sensor data [14]. Lee et al. expected Plexlines to aid in identifying children with developmental delay with its visualizations, but also identified the visualization's potential to act as an artifact for communication [17]. We base our study on this finding to further explore using behavioral visualization in clinical settings.

4.2 Visualizations for Communication

Previous work has covered using visualization for communicative purposes in various settings. Researchers have demonstrated that visualizations are effective in narrating a story [24], and how visualizations can be specifically designed to affect end-user interpretations [10]. Work by Hundhausen et al. also recognized the potential of visualizations in communicative context, and presented a preliminary characterization of six communicative dimensions found in visualizations [11]. In a similar vein, researchers have explored visualization enabled asynchronous collaboration. Heer et al. studied how like-minded people collaborate through discussion forums and visualization annotations [8], others have also explored synchronous collaboration in context of programming [16, 21] and tabletop visualizations [5, 12]. Previous work provides a rich description of how

the social aspect of visualization increased levels of engagement. Our work extends previous research in behavioral visualization tools and further investigate which specific features enable the visualizations to act as an effective means of communication between researchers, clinicians, and non-professionals.

5 METHODS

To compare and contrast Plexlines and EnGaze, we conducted a user study with researchers whose background is in child development and behavioral analysis. Half of the participants had worked in a clinical setting prior to starting research. We recruited participants by placing flyers in three local universities and through snowball sampling. Each participant spent approximately one hour for a session and received a $12 gift card.

The user study consisted of a preliminary interview, the Plexlines phase, the EnGaze phase, a semi-structured open-ended interview, and a comprehensive survey. During the preliminary interview, we asked for demographic information and professional experience related to children with developmental delays. Then, participants watched a five minute introductory video for Plexlines and explored the tool for around 15 minutes. This phase was paralleled for EnGaze with a five minute introductory video for EnGaze, followed by 15 minutes of exploration. The interface of the visualization tools can be seen in Figures 2 and 5. Both visualization tools provided 25 randomly selected RABC sessions for the study participants. Out of the 25 children, two were evaluated as at risk for ASD. To prevent order effects, seven participants started with Plexlines, and the other eight started with EnGaze.

After the exploration phase, we conducted a semi-structured open-ended interview. We first asked participants to compare and contrast the patterns they saw in Plexlines and EnGaze, then inquired about the strengths and weaknesses of each behavior visualization tool. We then asked for their thoughts on the representation of gaze in each visualization and which one would be more advantageous. We also inquired whether they could imagine using each visualization in their workflow, and upon an affirmative answer, we prompted for the imagined use cases of the visualizations. All interviews were recorded and transcribed for analysis. Then, we asked participants to tell us a story using Plexlines, EnGaze, or Both and to explain a child's behaviors to check the participants' comprehension of the tools. Participants concluded the session with a written survey that included Likert scale questions and open-ended questions comparing the two visualization tools in their features and use cases. The open-ended questions and the Likert scale questions are listed below. The five-point Likert scale ranged from poor (1) to excellent (5).

- Q1. Which visualization could serve better as a graphical longitudinal record of a child? Why?
- Q2. Can you imagine using such visualization of communicative behavior to teach clinicians or researchers about ASD?
- Q3. Which visualization is easier to read individually? Which visualization is easier to read in large groups in the webtool?

- L1. How well does [this visualization] show a child's response to examiner?
- L2. How well does [this visualization] identify children at risk?

Q1 and 2 were phrased with the intention of further exploring two prominent communication-related use cases of visualizations that were mentioned in previous work [15, 17]. In this study, we delved deeper to identify which specific features of each visualization enabled the above potential use cases. Lastly, we included Q3 as these tasks are often required in pre-screening sessions and to identify the function of EnGaze and Plexlines.

Kong et al. and Lee et al. [15, 17] asked questions similar to L1 and L2 to study EnGaze and Plexlines. Both visualizations were rated equally highly in their user studies. We hypothesized that a comparative study might elicit more critical responses since the participants had a chance to compare two visualizations side-by-side during the same session and could compare their usefulness for the given use case.

5.1 Participants

Our participants were clinicians and researchers, mostly female (13 out of 15). All of them had experience working with children with autism, except for P12 who had worked as a clinical consultant for a firm that makes augmentative and alternative communication devices for children with autism. Their mean years of experience in working with children with autism was 4.08 years (σ^2 = 3.42 years). Their majors in college varied from neuroscience, human biology, speech and language pathology to communication sciences and disorders, leading to diverse insights. Seven participants were graduate students and one was a senior in college. The majority of the participants had worked as a clinician (N=3), a therapist (N=3), a speech pathologist (N=2), or a special education teacher (N=2) prior to starting research in the domain. Although participants did not have prior experience with the RABC coding process, everyone was experienced in the general autism diagnosis process, either through performing diagnostic testing or analyzing and coding the sessions. Thus, they quickly became familiar with RABC based on their own experience, and successfully identified indicators of autism using the tools. Eleven participants, from three different universities, had a background in training parents for continuing ASD interventions at home, which led to insightful comments about using the tools in a collaborative setting.

6 RESULTS

According to the survey and interview results, participants clearly preferred Plexlines for comparison across sessions and EnGaze for visualizing a single session. This clear division in visualization preference for two distinct purposes is interesting considering that previous works [15, 17] have presented individual visualization as suitable for all these scenarios. Through a direct comparison of the two visualization tools, we were able to discover that Plexlines and EnGaze emphasized different granularities of behavior, which led to clear differences in the user study feedback. Plexlines focused on low granularity through emphasizing behaviors of longer duration while EnGaze presented high granularity by displaying behaviors of shorter duration, including micro-behaviors. Interview results furthermore uncovered the importance of showing and highlighting behavior patterns, an aspect not mentioned in previous studies.

6.1 Use Cases and the Preferred Visualization

Participants expressed a clear visualization preference in Q3 where fourteen out of fifteen study participants preferred EnGaze over Plexlines for visualizing an individual or a single session. On the other hand, thirteen participants preferred Plexlines for visualizing a group of sessions. This preference for single session tasks and multi-session tasks further led to their visualization preference for educating about ASD (Plexlines) and communicating with parents (EnGaze). All the use cases mentioned in this section stemmed directly from the participants' experience in working with children with autism and their parents. Within a few minutes of being introduced to Plexlines and EnGaze, participants got excited about the tools as they could clearly foresee how the visualizations would facilitate their routine clinical tasks, such as showing a summary of an autism intervention session to parents. Some participants (P1, P2, and P13) repeatedly asked about the availability of the tool as they wanted to use them in their practice.

EnGaze received higher scores than Plexlines for Likert questions L1 and L2, suggesting its usefulness in showing communicative behaviors of children. It received an mean score of 4.53 ($\sigma^2 = 0.51$) for showing responses (Plexlines $\mu = 3.73$, $\sigma^2 = 0.88$) and 4.13 ($\sigma^2 = 0.35$) for identifying children at risk (Plexlines $\mu = 3.8$, $\sigma^2 = 0.86$). There were no order effects.

6.1.1 EnGaze for Single Sessions and Communicating to Parents. EnGaze provided finer behavior granularity of a session through the detailed display of dyadic interactions and micro-behaviors. This higher granularity in EnGaze made clinicians feel it would be more effective for **communicating with parents** about their child and for parent training. "EnGaze might be better to use to explain to the parents, 'this was the baseline of your child's behavior. However, after the [outside] intervention, do you see there is a lot of [change in behavior.]' I think it'll be really useful for them, using EnGaze to show as supplemental intervention results" (P1). Visualizing higher granularity of behavior is useful because many communicative behavior milestones include micro-behaviors, which occurs in a fraction of a second. P12 provided one example of an important micro-behavior when s/he mentioned that EnGaze is especially useful for "working with a kid who has a developmental disability that affects social eye contact." Social eye contact is often a micro-behavior, and is easy to miss if the visualization solely focuses on lower granularity behaviors. Lack of eye contact is a sign of developmental delay and being able to track the frequency of eye contact can be extremely useful in tracking the developmental progress of a child. EnGaze appropriately supported a detailed examination of gaze by offering a separate page where gaze patterns are singled out and can be viewed in more detail (See Figure 5).

One participant suggested that visualizations would be useful for conveying these important but subtle child behaviors to parents, "I think this will also be a great tool for explaining to parents, a lot of times parents don't notice these things, they don't notice that their child isn't engaging or doesn't have eye contact that you can just compare to other children and say this is completely different" (P5). Once again, eye contacts are mentioned as an essential check point. Other essential micro-behaviors include looking back and forth between an object and a person for joint attention and short vocalizations as responses to examiner bids.

Eleven of our participants' research involved parent training, and thus they anticipated using EnGaze to make parents more self aware of their interaction with their child; "Plexlines is more child focused and [EnGaze] is more dyadic. It'd be more useful to show parents or to see how they're interacting [with the child]" (P10). More specifically, P9 mentioned that s/he would "mostly use [EnGaze] because eye contact with an examiner during the assessment is super important. We get most of our data from parent interactions, and we're looking both at the parent and the kid versus just the kid. Cause a lot of our research is parent intervention that we teach them how to do, so we're making sure they are doing it well, as well as if the kid is responding. Using the child-parent, it's so much more generalizable. You have the therapist and the speech pathologist come once a week but if the parent learns the principles and how to act, and in certain ways respond to that kid" (P9). S/he envisioned visualizations as a way in which parents could evaluate their own interaction with their children through visual feedback. After collecting a week or a month's worth of visualizations, parents could show moments of interest to the clinician, and clinicians could point out areas for improvement.

P12 conducted play therapy, for which s/he found EnGaze to be appropriate since it could highlight and provide context for the unanticipated micro-behaviors in **the naturalistic play session**. "You're looking at less structure. You're looking at how child communicates more independently on their own, outside of structured environments. [EnGaze provides the context] letting you know [...] from the kid's own perspective what were they needing to communicate, how do they communicate, what do they use to communicate" (P12). Because EnGaze presented all modalities and

micro-behaviors in one view, it was easier to see the exact behavior modality and timing that the child used to communicate. Assessing a child's current communication skills is important as it informs the clinicians of which communication skills to target for improvement.

6.1.2 Plexlines for Comparing Sessions and Educating about ASD. Thirteen participants preferred Plexlines for visualizing a group of sessions due to its low granularity, which led to a compact visualization (Q3). Although Plexlines contains details about examiner bids, participants perceived it as concisely providing an high level overview of a session because child and examiner behaviors across all modalities were visualized on one line. In contrast, EnGaze separated the child and examiner behaviors in five separate lines (Figure 8), providing a detailed low level view of the session. Participants (N=3) imagined "putting a bunch of EnGaze lines one after another [...] a bit overwhelming" (P12). The ease of comparing sessions led to two potential use cases for Plexlines: **educating people about ASD** and **showing structured protocol sessions**.

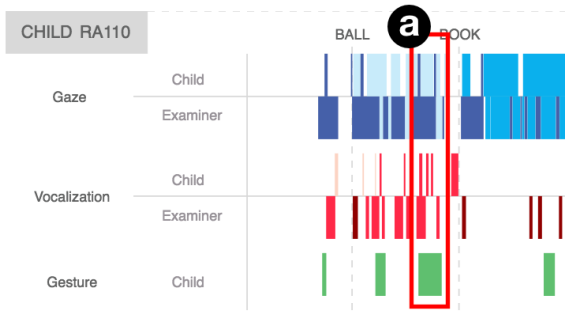

Fig. 8. EnGaze presents all three modalities (gaze, vocalization, and gesture) for the child and gaze and vocalization for the examiner. Examiner gesture data does not exist, resulting in a total of five lines of information. In (a), the child briefly glances at the examiner twice, shown as two slivers of navy rectangles on the very top row for gaze. At the same time, the child is pointing at an object, shown as a green rectangle in the very bottom row.

Educating people about ASD had been mentioned as a potential use case for visualizations "seeing that [the visualization] illustrates a lot of these patterns that we know as clinicians and researchers; but for the untrained eye, it would be difficult to take in through the video" (P11). Participants found Plexlines especially appropriate for this purpose because of its aggregate view feature (see Figure 4b) and its strength of visualizing multiple sessions. One participant (P1) imagined using the aggregate view as an ice breaker in an informational lecture on ASD to compare the behavioral characteristics of children with autism with those of typically developing children. She said "if we're presenting something at a conference, we don't want to just compare two participants, one with autism and one without autism, and here's the difference. We want to have enough data to see that this is the average range of [an] individual with autism, and this is the average range of [an] individual without autism. And here is a range of communication skills that they're using." She could imagine using one aggregate view for each group of children or having two columns of Plexlines – one column for each group with six to seven children per group – since multiple Plexlines easily fit into one page. P12 mentioned that Plexlines would be suitable for **a structured environment** since a user can organize behaviors through compare sessions by the phases in a protocol (see Figure 9). Different reactions to the same phase of a protocol could easily be compared when the sessions are aligned by phases. Focusing on longer behaviors such as in Plexlines could also provide "[a] more useful scoring measure for kids you know have decent social communication,

but use for different type of diagnosis. [The child could] potentially have a speech disorder or different types of anxiety issues that wouldn't necessarily affect social eye contact, social gaze, social interaction" (P12).

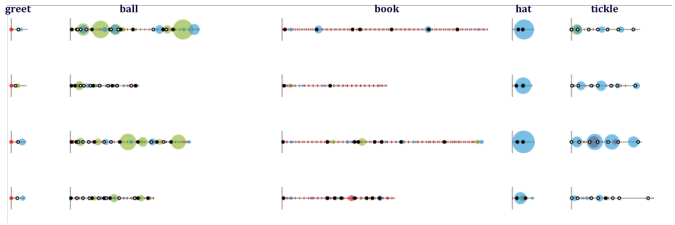

Fig. 9. Four Plexlines divided based on the five different phases. This view allows users to compare individual phases of the RABC instead of the entire session.

6.1.3 Balanced Preference for a Longitudinal Record. Unlike Q3 and the Likert scales questions where one visualization was clearly preferred over the other, responses were split for which visualization could serve as a graphical longitudinal record of a child (Q1); eight participants selected EnGaze and six selected Plexlines. P15 commented on how s/he envisioned using the visualization as a longitudinal record of a child to describe child development to a parent. "I could see [Plexlines and EnGaze] being helpful when they do come in for an assessment having this data at certain time points. So I could be like 'they are really making improvements with their gaze, but their gesture is not [improving].' " Participants who preferred EnGaze preferred it for the visualization of detailed behaviors, ease of perceiving the duration of each action, and the ability to highlight moments of joint attention. Participants who chose Plexlines mentioned the simple nature of Plexlines and ease of comparing multiple visualizations as their reason for selecting Plexlines.

6.2 Supporting Identification of Behavior Patterns

Rather than using individual behaviors (e.g., a single eye contact or utterance of a word) as a basis of analysis, our participants mainly used behavior patterns, such as engaging in joint attention or pointing while talking, to describe children in both Plexlines and EnGaze. Participants searched for behavioral patterns because the contextual information around a behavior influences the importance and implications of the behavior. For example, contextual information determines whether a child behavior was elicited by an adult (i.e. a response) or unprompted (i.e. an initialization). As one of the participant's research focused on teaching responses and initializations in an autism intervention, s/he (P2) found this distinction of behaviors based on context essential. S/he saw that this visualization could be useful for pointing out missed opportunities to parents where they could have waited longer for a response from their child.

As EnGaze had been specifically designed to identify moments of joint attention, participants often mentioned joint attention related patterns. "I like seeing the examiner's behavior and eye contact and all of that. I think that's really useful. The turn taking, [and] the joint attention is also really nice" (P10). However, participants independently identified other common behavior patterns of a child through the visualization. In general, these behavior patterns were closely related to common developmental milestones of children such as gaze alternation and turn taking. Similarly, participants identified behaviors patterns in Plexlines as well. Behavior patterns found in Plexlines during the study include *examiner bid response pattern* as Plexlines specifically emphasized examiner's bids in RABC by representing them as black dots (see Figure 3b). "[Plexlines] is a good way to explain the interactions between child and examiner and to determine how the child

responds. It allows for a quick way to identify possible red flags, [such as] a lot of examiner bids and not many child responses" (P5).

The above examples show the importance of designing a behavior visualization such that it not only conveys the frequency of individual behaviors but provides the contextual information and behavior patterns. One participant (P1) elaborated how providing contextual information would improve her current practice, which involves two separate graphs – a line graph to show the percentage of opportunities (i.e., the mother's strategy use) to which the child responded and a bar chart to show the number of times the child initiates communication (See Figure 10) – similar to the figure found in [20]. Because the response and initialization is taken out of context, s/he had to revisit the videos to ensure their implications. As Plexlines and EnGaze both display both the context and the frequency of child behavior in one visualization, s/he was able to see the progression of a session more clearly. S/he stated that the ability to see the whole session in a single visualization made the visualization useful for explaining a session to a parent.

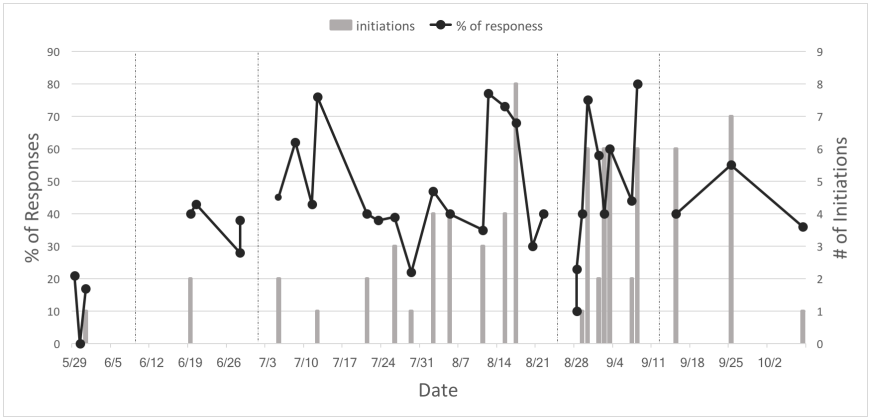

Fig. 10. A remake of a visualization that one of our participant currently uses for their analysis of communication behavior session results. The line depicts the responses of a child for all engagement opportunities in percentages, and the bar chart is the number of times the child initiates communication.

7 DISCUSSION

The user study with Plexlines and EnGaze revealed that a comparative study can identify features and differences that are not brought out through a single-tool study, and that presenting and supporting different behavior granularity is linked to the function of the tool. Specifically, this study contextualizes Shneiderman's mantra of "Overview first, zoom and filter, then details-on-demand" [25] to the domain of behavioral data where low granularity visualization such as Plexlines can provide an overview and high granularity visualization such as EnGaze can provide details-on-demand. We illustrate how three granularity-related features — displaying micro-behaviors, the arrangement of the three different modalities, and coordinating dyadic interactions — can be used for presentation of behavioral data for autism pre-screening. Then, we give examples for how each of the features can be extended to health care settings beyond the field of ASD.

7.1 Displaying micro-behaviors

Participants found visualizing micro-behaviors important as it is an essential part of communication (e.g., quick alternating glances in joint attention). The different emphasis on macro and micro behaviors stemmed from the shapes used in the visualizations. The impact of design decision of

shapes went beyond the aesthetics of the visualization. Plexlines uses circles to visualize behavior, whereas EnGaze uses rectangles. Because circle width is directly proportional to the circle height, the duration of a behavior is represented in two dimensions in Plexlines (i.e., behaviors of longer duration are taller as well as wider) while only in one dimension in EnGaze (i.e., behaviors are of the same height regardless of their duration) as shown in Figures 3 and 6. As a result, micro-behaviors have little spatial presence in Plexlines, as larger circles can easily obscure smaller circles, rendering them invisible to the naked eyes of its viewers. The same micro-behaviors would be clearly visible in EnGaze. One participant stated that viewing circles helped her easily compare across sessions since "you can clearly see circles grow over time" (P4). However, other participants (N=3) criticized that estimating the area or the radius of a circle in Plexlines is difficult and found reading the bar width more intuitive in EnGaze. P2 pointed out that the circular representation also exaggerated the duration of a behavior.

Participants generally found micro-behaviors informative, but P2 reported that the rapidly alternating bars in EnGaze "make my eyes get tired because there are too many different bars with too many different colors. It's more complex than Plexlines' circles." This suggests that the inclusion of micro-behaviors could degrade the viewing experience because of the resulting dense and complex visualization. We also infer that a high granularity visualization with an emphasis on behaviors of longer duration would be useful for situations where the duration of a behavior is an essential part of analysis, such as in an Attention Deficit Hyperactivity Disorder (ADHD) intervention where the goal is to increase the duration of the child's focused attention [3]. Considering these benefits and weaknesses of displaying and emphasizing behaviors of different granularity, providing multiple levels of detail in communicative behavioral visualization is helpful.

7.2 Arrangement of Modalities

The arrangement of the three behavioral modalities influenced the emphasis of certain behaviors. Gaze, vocalization, and gesture are all vertically aligned on a line in Plexlines (See Figure 3). Although bigger circles are algorithmically placed behind the smaller circles to avoid obscuring concurrent behaviors, longer behaviors can still downplay shorter co-occuring behaviors as the circles overlap. The co-occuring behaviors were still visible in Plexlines, but could not convey detailed information. On the other hand, EnGaze offered the option of displaying each modality on a separate adjacent line, and provides a equal level of presence to co-occuring behaviors of different durations (See Figure 8). Thus, micro-behaviors (e.g., a quick glance) can be observed even if there is a longer co-occuring behavior in another modality (e.g., pointing to an object) shown as thin dark blue lines in Figure 8a.

Participants valued being able to see all the modalities in adjacent rows, "What we do a lot of in coding is [...] matching two categories, combined communication like eye contact with speech, eye contact with gesture, or gesture with speech. That would definitely be the number one thing to look for in our projects right now" (P9). This use case is not limited to child-clinician interactions for ASD screening. Another clinical setting where visualization of multiple modalities would be useful is treatment for social anxiety disorder. Social anxiety disorder can result in atypical patterns in gaze (e.g., shorter eye contact), speech (e.g., longer silent pauses, stuttering), and gestures (e.g., fewer smiles or nods) [4, 9]. The symptoms vary by person and different combinations of modality could be affected. For similar scenarios where viewing multiple modalities is essential, we propose visualizing all modalities on a single line for overview and separating the three modalities into individual lines upon zooming in.

7.3 Coordinating dyadic interactions

Currently, many established works in autism diagnosis focus solely on the child behavior [18, 19] and is accepted as the norm. However, participants highly praised the inclusion of examiner behavior in our study, and also in previous Kong et al.'s study on EnGaze [15]. Providing a symmetrical behavior view of both parties in a dyad allows the viewers to understand the context of the dyadic interaction, which is essential in autism pre-screening. We propose an overview that focuses on one person's behaviors, and expand to dyadic behaviors upon zoom until both parties have equal presence. We further suggest allowing users to choose the person to focus on in the overview.

The study of dyadic interactions is not limited to autism diagnosis. People who were not part of our study showed interest in the tool for its ability to show dyadic interaction. One person from a family resiliency center proposed that the tool could be used to visualize parent-child interactions during dinner time. The visualization could depict the frequency and the quality of engagement through depicting who is initiating the conversations and how the other person is responding. Another therapist wanted to experiment with the tool for therapist-patient interactions. A local hospital also requested access to the tool for their clinicians. This display of interest from people outside of our study shows that visualizing dyadic interaction extends beyond ASD-related scenarios.

8 FUTURE WORK AND LIMITATIONS

Our work presents a study of visualizations for clinical behavioral dyadic communication. Although the participants had expertise in the autism prescreening process, not all of them had clinical experience. Further verification from an expanded pool of clinicians is required to confirm participants' statements and suggested uses. In addition, while our participants envisioned using the visualizations when talking to parents, we lack parental feedback in our study. Previous studies have shown that parents express frustration over the lack of information provided during the diagnostic process [22]. Although clinicians envisioned that visualizations would facilitate communication with parents, one participant expressed concern that singling out their child in a visualization might agitate parents, "I don't know if all parents would exactly like that, it could be little offensive..." (P9). The deployment of the tools in a collaborative setting involving clinicians, parents, and children would help concretize the findings of this paper by adding the parental perspective and substantiating the clinicians' perspective.

We compared two coordinated communication behavioral visualizations, specifically designed around an autism screening instrument. One can imagine other visualizations built around this instrument. Although these visualizations evolved over many iterations to satisfy their goals, it is possible that our results are influenced by our selection of these specific visualizations. More work is necessary to show generalizability to all coordinated communication behavior visualizations in ASD and other communication domains.

9 CONCLUSION

In this paper, we presented a comparative study of two visualizations of communicative behavior designed for ASD pre-screening. Our study results show that a qualitative comparison of multiple visualizations can bring out insights that are not found when each visualization is studied individually. Participants preferred EnGaze for viewing an individual session, and Plexlines for viewing a group of sessions. This visualization preference arose from different behavior granularity emphasis in the two visualizations. Prior studies that presented EnGaze and Plexlines individually did not identify this difference. The survey and interview results further revealed showing and highlighting behavior patterns as a favorable feature in behavior visualization. Based on the findings that

granularity of visualizations led to envisioned use cases, we propose three key granularity-related features – separating modalities, coordinating dyadic interactions, and displaying micro-behaviors – and present potential use cases beyond ASD screening.

REFERENCES

[1] Tony Bergstrom and Karrie Karahalios. 2007. Seeing more: visualizing audio cues. *Human-Computer Interaction–INTERACT 2007* (2007), 29–42. https://doi.org/10.1007/978-3-540-74800-7_3

[2] Tanja Blascheck, Kuno Kurzhals, Michael Raschke, Michael Burch, Daniel Weiskopf, and Thomas Ertl. 2014. State-of-the-art of visualization for eye tracking data. In *Proceedings of EuroVis*, Vol. 2014. https://doi.org/10.2312/eurovisstar.20141173

[3] Norbert Börger and Jaap van der Meere. 2000. Visual behaviour of ADHD children during an attention test: an almost forgotten variable. *Journal of child Psychology and Psychiatry* 41, 4 (2000), 525–532. https://doi.org/10.1111/1469-7610.00637

[4] Sandra Daly. 1978. Behavioural correlates of social anxiety. *British Journal of Clinical Psychology* 17, 2 (1978), 117–120. https://doi.org/10.1111/j.2044-8260.1978.tb00252.x

[5] Aaron Genest and Carl Gutwin. 2012. Evaluating the effectiveness of height visualizations for improving gestural communication at distributed tabletops. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. ACM, 519–528. https://doi.org/10.1207/s15327051hci1903_3

[6] Joshua Hailpern, Karrie Karahalios, Laura DeThorne, and Jim Halle. 2010. Vocsyl: Visualizing syllable production for children with ASD and speech delays. In *Proceedings of the 12th international ACM SIGACCESS conference on Computers and accessibility*. ACM, 297–298. https://doi.org/10.1145/1878803.1878879

[7] Yi Han, Agata Rozga, Nevena Dimitrova, Gregory D Abowd, and John Stasko. 2015. Visual analysis of proximal temporal relationships of social and communicative behaviors. In *Computer Graphics Forum*, Vol. 34. Wiley Online Library, 51–60. https://doi.org/10.1111/cgf.12617

[8] Jeffrey Heer, Fernanda B Viégas, and Martin Wattenberg. 2007. Voyagers and voyeurs: supporting asynchronous collaborative information visualization. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 1029–1038. https://doi.org/10.1145/1240624.1240781

[9] Stefan G Hofmann, Alexander L Gerlach, Amy Wender, and Walton T Roth. 1997. Speech disturbances and gaze behavior during public speaking in subtypes of social phobia. *Journal of Anxiety Disorders* 11, 6 (1997), 573–585. https://doi.org/10.1016/S0887-6185(97)00040-6

[10] Jessica Hullman and Nick Diakopoulos. 2011. Visualization rhetoric: Framing effects in narrative visualization. *IEEE transactions on visualization and computer graphics* 17, 12 (2011), 2231–2240. https://doi.org/10.1109/TVCG.2011.255

[11] Christopher D Hundhausen. 2005. Using end-user visualization environments to mediate conversations: a 'Communicative Dimensions' framework. *Journal of Visual Languages & Computing* 16, 3 (2005), 153–185. https://doi.org/10.1016/j.jvlc.2004.11.002

[12] Petra Isenberg, Danyel Fisher, Sharoda A Paul, Meredith Ringel Morris, Kori Inkpen, and Mary Czerwinski. 2012. Co-located collaborative visual analytics around a tabletop display. *IEEE Transactions on visualization and Computer Graphics* 18, 5 (2012), 689–702. https://doi.org/10.1109/TVCG.2011.287

[13] Julie A Kientz, Sebastian Boring, Gregory D Abowd, and Gillian R Hayes. 2005. Abaris: Evaluating automated capture applied to structured autism interventions. In *International Conference on Ubiquitous Computing*. Springer, 323–339. https://doi.org/10.1007/11551201_19

[14] Jennifer Kim, Melinda Snodgrass, Mary Pietrowicz, Karrie Karahalios, and Jim Halle. 2013. BEDA: Visual analytics for behavioral and physiological data. In *Workshop on Visual Analytics in Healthcare. Washington DC*. 23–27.

[15] Ha Kyung Kong, John Lee, Jie Ding, and Karrie Karahalios. 2016. EnGaze: Designing Behavior Visualizations with and for Behavioral Scientists. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*. ACM, 1185–1196. https://doi.org/10.1145/2901790.2901870

[16] Mikko-Jussi Laakso, Niko Myller, and Ari Korhonen. 2009. Comparing Learning Performance of Students Using Algorithm Visualizations Collaboratively on Different Engagement Levels. *Educational Technology & Society* 12, 2 (2009), 267–282. https://doi.org/10.1145/1268784.1268857

[17] John Lee, Ha-Kyung Kong, Sanny Lin, and Karrie Karahalios. 2016. Plexlines: Tracking Socio-communicative Behaviors Using Timeline Visualizations. In *AMIA Annual Symposium Proceedings*, Vol. 2016. American Medical Informatics Association, 1890.

[18] Catherine Lord, Susan Risi, Linda Lambrecht, Edwin H Cook, Bennett L Leventhal, Pamela C DiLavore, Andrew Pickles, and Michael Rutter. 2000. The Autism Diagnostic Observation Schedule–Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of autism and developmental disorders* 30, 3 (2000), 205–223. https://doi.org/10.1023/A:1005592401947

[19] Catherine Lord, Michael Rutter, and Ann Couteur. 1994. Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of autism and developmental disorders* 24, 5 (1994), 659–685. https://doi.org/10.1007/BF02172145

[20] Hedda Meadan, Melinda R Snodgrass, Lori E Meyer, Kim W Fisher, Moon Y Chung, and James W Halle. 2016. Internet-based parent-implemented intervention for young children with autism: A pilot study. *Journal of Early Intervention* 38, 1 (2016), 3–23. https://doi.org/10.1177/1053815116630327

[21] Niko Myller, Roman Bednarik, Erkki Sutinen, and Mordechai Ben-Ari. 2009. Extending the engagement taxonomy: Software visualization and collaborative learning. *ACM Transactions on Computing Education (TOCE)* 9, 1 (2009), 7. https://doi.org/10.1145/1513593.1513600

[22] Lisa A Osborne and Phil Reed. 2008. Parents' perceptions of communication with professionals during the diagnosis of autism. *Autism* 12, 3 (2008), 309–324. https://doi.org/10.1177/1362361307089517

[23] Opal Ousley, Rosa Arriaga, Gregory Abowd, and Michael Morrier. Jan 2012. Rapid Assessment of social-communicative abilities in infants at risk of autism. *Technical Report CBI-100, Center for Behavior Imaging, Georgia Tech* (Jan 2012).

[24] Edward Segel and Jeffrey Heer. 2010. Narrative visualization: Telling stories with data. *IEEE transactions on visualization and computer graphics* 16, 6 (2010), 1139–1148. https://doi.org/10.1109/TVCG.2010.179

[25] Ben Shneiderman, Catherine Plaisant, and Bradford W Hesse. 2013. Improving healthcare with interactive visualization. *Computer* 46, 5 (2013), 58–66. https://doi.org/10.1109/MC.2013.38

[26] Maryam Tohidi, William Buxton, Ronald Baecker, and Abigail Sellen. 2006. Getting the right design and the design right. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*. ACM, 1243–1252. https://doi.org/10.1145/1124772.1124960

Received April 2017; revised July 2017; accepted November 2017