Computer Supported Cooperative Work (CSCW) (2021) 30:393–424
https://doi.org/10.1007/s10606-021-09402-7

Tensions in Representing Behavioral Data in an Electronic Health Record

Gabriela Marcu[1,*], Anind K. Dey[2] & Sara Kiesler[3]
[1] *University of Michigan, Ann Arbor, MI, USA (E-mail: gmarcu@umich.edu);* [2] *University of Washington, Seattle, WA, USA;* [3] *Carnegie Mellon University, Pittsburgh, PA, USA*

Accepted: 9 May 2021

Abstract. Taking an action research approach, we engaged in fieldwork with school-based behavioral health care teams to: observe record keeping practices, design and deploy a prototype system addressing key challenges, and reflect on its use. We describe the challenges of capturing behavioral data using both paper and electronic records. Creating records of behaviors requires direct observation, and as a result the record keeping responsibility is challenging to distribute across a care team. Behavioral data on paper must be transferred and prepared for reporting, both inside the organization and to stakeholders outside of the organization. In prototyping a computerized working record, we targeted user needs for capturing details of a behavioral incident in the moment. Challenges persisted through the transition from paper to our prototype, and based on these empirical findings over two years of fieldwork, we present five tensions in representing behavioral data in an electronic health record. These tensions reflect the differences between entering behavioral data into the record for intraorganizational use versus interorganizational use.

Key words: Workflow, Interprofessional collaboration, Documentation, Health records, Action research

1. Introduction

Health records come in many different forms. More formal conceptualizations such as medical charts or electronic health records help to standardize practices of health care teams, while informal practices of jotting down notes on paper or exchanging information support the ad hoc aspects of team work. Fitzpatrick (2004) uses the term 'working record' to differentiate from a passive information repository. A health record 'at work' takes various forms because it is maintained and evolves through practical aspects of delivering health services. For example, components of the working record can take varying forms when authored by different members of the team, and for different intended audiences, lifespans, locations, or purposes (Fitzpatrick 2004). Health records can enforce standardization of practices, but they can also be designed to provide flexibility and adaptivity for local and immediate needs of teams (Fitzpatrick 2004; Berg 2004).

One practical local need is managing subjective data, which are not easily standardized. Subjective aspects of patient care and coordination are just as important as objective ones (Erikson 1958), but they are difficult to represent in the record (Mentis 2010). To explore these aspects, we focus on behavioral health, describing local needs and demands through fieldwork studying care teams supporting children with behavioral challenges. The individualized nature of these services involve a subjective process of identifying behaviors of concern, then operationalizing, measuring, and monitoring them (Marcu et al. 2013, 2016). Children's behaviors are recorded manually based on direct observation, and monitored to determine the effectiveness of interventions.

Over two and a half years, we studied this process among care teams by examining their evolving use of the working record in practice. Our focus on the role of records among teams providing psychiatric care for children represents an understudied context. Building on limited work in this area, we examine the role of schools within the ecosystem supporting children's behavioral health. Complementing Saario et al. (2012), who studied coordination across organizations, we focus on the intraorganizational practices of record keeping by school-based teams. These care teams monitor and address behaviors that affect a child's ability to learn and develop interpersonal relationships. Early intervention is most effective, and schools play an important role in prevention and intervention (Greenwood et al. 2008). Under U.S. law, schools are required to identify needs and provide services for children that can be classified as having behavioral disabilities (Yell et al. 2018). Children with autism spectrum disorders, conduct disorders, anxiety disorders, neurological impairment, emotional impairment, and other types of diagnoses can exhibit disruptive behaviors that may lead to services within a school setting. School-based behavioral health teams include such professions as special educators, school psychologists, psychiatrists, behavior analysts, therapists, and social workers.

Behavioral health services for children have a rich set of dependencies across stakeholders and settings in a child's daily life—especially between healthcare and education—which has made them a focus of sociotechnical research (Amir et al. 2015; Hayes et al. 2010; Marcu et al. 2014). Care coordination is complex because of the necessity to navigate across different systems of care (Council on Children with Disabilities 1999). Coordination is 'the process of managing dependencies among activities,' and what it looks like varies across settings (Malone and Crowston 1994). Behavioral health involves what Kaziunas et al. (2019) call ecologies of care, across the formal healthcare system and community services such as those provided by faith-based groups. Their ethnographic account describes how coordination across an ecology of care could help individuals manage their health, but instead there tends to be fragmentation. Behavioral health is additionally dependent on managing one's health at home (Barbarin et al. 2015), and this is particularly important for children and others who have caregivers playing an active role in their care. Children's behavioral health involves a range

of stakeholders, from parents and other caregivers in the home, to school staff responsible for their care during the day, and behavioral specialists in other community and clinical settings. When children exhibit behaviors that are concerning, these stakeholders must coordinate across organizational and professional boundaries. Coordination is so interdependent and challenging that there can be tensions and even breakdowns, as our prior work showed at the boundary between home and school (Marcu et al. 2019).

In this paper, our focus on the role of behavioral records for coordination shows how those who were responsible for creating and maintaining records for behavioral monitoring struggled with the local demands of this work, and wanted technological supports. We took an action research approach to investigate their paper-based record keeping practices, design a novel electronic health record, and conduct a deployment study of our resulting prototype. Throughout these stages of our research, we uncovered tensions at the team and organizational levels around how behavioral data should be collected, represented, and used. Accordingly, this paper addresses the following research question: *What are the tensions involved in representing behavioral data in an electronic health record, for use in care coordination?* We contribute rich descriptions of practices with behavioral data, and a study of technology-in-use, explaining how tensions in representing behavioral data in an electronic health record have to do with balancing intra- and inter- organizational information needs. Our study of these teams addresses the paucity of literature focused on allied health professionals (Fitzpatrick and Ellingsen 2013).

2. Related work

2.1. Record keeping and care coordination over time

For the purposes of making informed decisions about care coordination over time, members of a behavioral health care team are reciprocally interdependent (Marcu et al. 2014). Reciprocal interdependence exists in a relationship in which both parties rely on one another for data to be able to complete their work (Thompson 1967). Reciprocal interdependence requires more than data sharing, it involves frequent communication and knowledge sharing. Children's behavioral health care teams are reciprocally interdependent for several reasons: data are collected by multiple team members and require aggregation; team members have unique experiences and data about one child; and team members are drawing from knowledge across different areas of expertise (Marcu et al. 2014). Team members work with the same child in different contexts, so they each develop unique knowledge based on the behaviors they witness within different situations. Care teams integrate these perspectives in understanding a child's progress, informing care decisions over time (Marcu et al. 2016). In this process, records can be a helpful tool for care teams to construct an unfolding narrative as a way of reasoning about a case (Mønsted et al. 2011).

In their study of teams caring for children with complex medical needs, Amir et al. (2015) describe the characteristics of complex care that make it difficult for a team to coordinate. For example, care plans require consensus across the team, there is continual distributed revision of plans, and different providers operate on syncopated time scales. Amir et al. therefore explain how technology can support care teams by enabling them to coordinate around a shared care plan that is 'ever present,' efficient, and adaptable for varying activities. Care teams in behavioral health experience similar challenges, with complexity of care arising in large part from the difficulty of measuring and monitoring progress—that is, behaviors are subjective and difficult to measure. Yet record keeping in the context of behavioral health has been understudied. The development and implementation of behavioral health records can be challenging due fragmentation among organizations (Kaziunas et al. 2019). Some of the key behavioral data that could facilitate coordination are not captured in electronic health records, but communicated anecdotally or via informal record keeping (Marcu et al. 2016; Saario et al. 2012).

Thus, in order to apply recommendations such as those from Amir et al. to design in behavioral health, we must first understand how to support care teams to effectively record behavioral data. In this study of care teams' work with behavioral records, we draw on the concept of local universality, which explains how the standardization of health records 'always rests on real-time work, and emerges from localized processes of negotiations and pre-existing institutional, infrastructural, and material relations' (Timmermans and Berg 1997). In light of the difficulty of standardizing behavioral records, we are interested in how these negotiations unfold with behavioral data.

2.2. Record keeping on paper

A body of literature has demonstrated the prevalence of record keeping on paper, largely due to its flexibility in supporting clinical practices. For example, fieldwork has investigated the use of paper records (Park et al. 2013; Jagannath et al. 2018), multiple texts (Christensen 2016) and other non-digital artifacts (Bardram and Bossen 2005) to support hospital work. Based on a study of paperwork in a hospital, Østerlund argues that technology needs to support 'continuity, improvisation, and change' embedded in practices with documents (2008). Park, Pine, and Chen describe how electronic medical records in hospitals do not support informal documentation practices required as part of clinical work (2013). Ellingsen and Monteiro (2003) describe various forms of knowledge representations in large hospitals, including paper and electronic records. These knowledge representations are enacted fluidly over time through texts and narratives. Records generated are rendered as credible knowledge through work within the team, and they 'perform essential tasks in regulating, coordinating and controlling the organisation of work both within and outside the hospital' (p. 226).

Many studies of record keeping among multidisciplinary teams have focused on hospital work, and included professions such as: nurse, social worker, physiotherapist, pharmacist (Fitzpatrick 2004); and physician, radiologist, pathologist (Kane and Luz 2011). We complement this literature by studying care teams providing behavioral health services for children outside of the hospital setting (e.g., school, community). Work in this area has been more limited, but includes Saario et al.'s (2012) investigation of documentation on child health and welfare, and how it supports communication across organizations and professions. Their study found that tacit knowledge was not passed along in electronic documentation, which only contained factual information, whereas paper tended to include more details as the record keepers took notes in the process of trying to make sense of a child's case. Similarly, our prior work has discussed the persistence of paper records among multidisciplinary care teams for children with special needs (Marcu et al. 2013).

2.3. Tensions with computer-assisted behavioral health monitoring

Given the challenges of capturing behavioral data over time, and reviewing potentially large amounts of rich data in order to monitor progress, automation and other supports have become a focus of computing research. Computer-assisted monitoring includes automated data capture via audio, video, or sensors; and it also includes manual data entry into a specially designed computer interface instead of paper.

Hayes et al. (2004) combined automation and manual data entry in developing capture applications for recording of behavioral data during one-on-one therapy sessions at school. They integrated behavioral data handwritten on their tablet interface with automatically captured video recordings of the behavior, to help care teams monitor progress across time and place. Hayes and Abowd (2006) reported that 'simultaneous data gathering' in school while working with children with behavioral needs was viewed by their participants as too difficult and unrealistic to manage. All of their participants wanted 'to record and to share rich media including video' as a way of addressing challenges with data collection. Consequently, the tensions explored in their work had to do with the associated tradeoffs of privacy, control, and surveillance.

Despite the promise of capture and access technologies, a decade and a half after this work was done, paper and pencil continue to be a common form of behavioral data collection. Even electronic health records, which are common elsewhere in healthcare services, remain difficult to implement in behavioral health (Kaziunas et al. 2019). We therefore complement this literature by focusing on manual rather than automated capture of behavioral data, and perform a sociotechnical analysis identifying a different set of tensions, which accompany computer-assisted manual entry of behavioral data.

In past work, we described how care teams struggled, even with computer-assisted data collection, to manage behavioral records as part of their work (Marcu

et al. 2019). We found that in order to be useful for reflection, behavioral data need to be captured with four qualities—reliability, meaningfulness, granularity, and availability—all of which are difficult to maintain. We have also attributed such challenges with behavioral data in part to the diffusion of responsibilities for care and coordination, an aspect of teamwork in behavioral health that is organic and difficult to design for (Marcu et al. 2016). In the present paper, therefore, we describe the tensions present in a team and organization as they work to collaboratively operationalize, measure, and monitor behavioral data in the electronic health record. In behavioral health, where standardization is not possible and team members are interdependent for reflecting on the data before they can act on it, we need a better understanding of tensions and how they arise around behavioral data.

3. Setting and methods

Figure 1 illustrates the methods used in this study, which unfolded over three phases. In the following section, we describe the setting in which this study took place. Then, we describe how formative fieldwork began and continued throughout the study with ongoing data analysis, as we designed a prototype and studied its use in a deployment.

3.1. School-based behavioral health services setting

The setting for this study was a school-based partial hospitalization program, providing behavioral and mental health services at elementary grade levels. Integrating treatment typical of psychiatric clinics within the school day, this type of program enables a child to live at home and attend school while receiving support to address behavioral, emotional, or social challenges on a daily basis. School districts will recommend placement in a partial hospitalization program when a child has had significant behavioral difficulties in a regular school. U.S. law guarantees 'free and appropriate public education' to all children with disabilities, and special education eligibility includes behavioral and emotional disabilities which

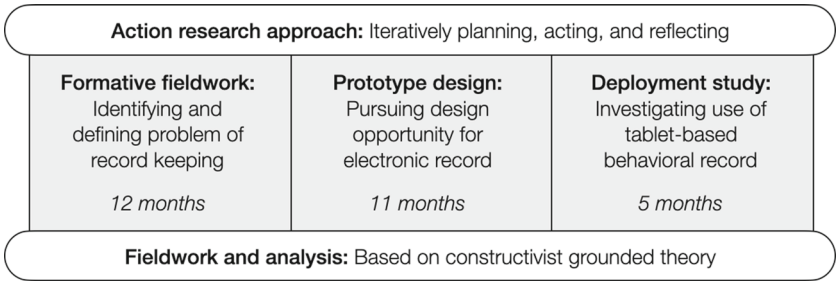

Figure 1. Overview of methods used over the three phases of this work.

require mental health services (Yell et al. 2018). Children may be recommended for this type of special education placement when they have diagnoses such as autism spectrum disorders, neurological impairments, or anxiety disorders. Children were grouped into classes based on their behavioral needs, as some children were functioning one or two grade levels behind what would be typical for their age.

Within an ecosystem of varying placements across a school district, this program's role was to provide support to students in the form of behavior management, cognitive behavioral therapy, and potentially psychiatric medication. Educational placement for each child in the school district was reviewed regularly alongside treatment decisions. For example, when a child's behavioral needs surpass the services available at this program, they may be referred to a psychiatric hospital for inpatient treatment. Some children make enough progress behaviorally that are transition to a different placement in the school district for gradual reduction of behavioral supports, which enables them to gain more independence. Such a transition also results in a less costly placement for the school district, for example with a special education classroom within a typical school.

3.2. Formative fieldwork and data analysis

Our fieldwork in the program began by exploring use and non-use of technology more broadly, especially with regard to delivery of services, before identifying the challenges of record keeping as a key concern. We used an action research approach, which values the local perspective and expertise of the participant by treating them as a partner in developing then acting on research goals. (Hayes 2011). We found that care teams struggled more with incorporating technology they felt was needed for managing records, than they did in regards to any technologies that might support delivering care to children.

In defining the problem of record keeping together, and then beginning to discuss a potential prototype, we engaged in 12 months of observation and interviews with the diverse members of care teams providing services: psychiatrists, mental health therapists, behavioral specialists, educators, paraprofessional educators, and clinical supervisors. We observed them in classrooms as they worked with children, in team meetings as they reviewed data to discuss cases, and during informal interactions in school hallways. While engaging with care teams to learn how they record and use behavioral data, we took a design anthropology perspective focused on the meaning of data, interactions, and artifacts involved in their team work (Clarke 2011; Gunn et al. 2013).

Our approach to collecting and analyzing field data was based on constructivist grounded theory (Charmaz 2006). We focused on gathering rich data using theoretical sampling across team members and contexts. Open coding was used to identify challenges experienced by members of the care team and describe their needs. Axial coding was then used as fieldwork continued, with constant

comparison of data previously collected, in order to develop themes. Theoretical sufficiency was reached when themes were representative of diverse team members' experiences, and also represented our collective understanding about the design opportunity. The themes then guided the main functionality of a prototype. Fieldwork continued throughout the deployment, and we used ongoing data collection and analysis to keep refining themes. We also compared themes to the literature as they began to lead us to our main insights about the tensions involved in capturing and representing behavioral data in electronic health records.

3.3. Prototype design and deployment study

Our action research process entered a new phase when an opportunity for design arose, and with our partners in the school we agreed to take action through design and deployment of a novel electronic health record. The first author then led a team of nine researchers in ongoing fieldwork to inform the design of a prototype. On a weekly basis, these researchers would discuss themes emerging from the field and iterate on concepts for the prototype through rapid prototyping. Ideation and prototyping were then validated with care teams during fieldwork. Input on mockups throughout the 11-month design process was primarily provided by those team members responsible for record keeping. We continued to iterate on the design of the prototype, which we eventually named Lilypad, until it was ready for use. Care teams outlined the core functionality they would need in order to begin using Lilypad in their work, which we used to determine together when the prototype was ready for deployment.

We modeled a field deployment (Siek et al. 2014) on in-the-wild studies (Bonsignore et al. 2013; Johnson et al. 2012; Marshall et al. 2011; Messeter and Molenaar 2012), to examine the use and impacts of the technology in situ, with real world social scenarios and constraints. To this end, Lilypad was deployed in three classrooms within the same school, serving children from second to sixth grades. Each classroom was staffed by four to five practitioners working with four to nine children. Ongoing fieldwork during Lilypad use constituted observation within the physical classrooms, as well as shadowing team members across different locations within the school building, and observing care team meetings. Observation across these different locations and activities enabled us to understand the coordination of a child's care team, which was made up of individuals across the organization (e.g., classroom staff, therapeutic staff, psychiatrists, and clinical supervisors). Fieldwork focused on seven teams each coordinating care for one child, comprised of 24 total care team members.

The deployment study lasted 24 weeks, during which time we conducted 76 hours of participant observation in 48 total visits, averaging 4.5 hours per week. 24 participants were recruited from among the school staff we spoke with during prior research stages, and they were compensated $20 per week during the deployment. The features of Lilypad were rolled out gradually, to ease their transition

to using it. Most fieldwork hours were spent with record keepers, who used the Lilypad system most frequently as the gatekeepers of the data. We also observed 27 care team meetings with permission from parents.

4. Record keeping practices and artifacts

Records served to document instances of behavior throughout the school day, using the frequency or duration of a certain type of behavior. Record keeping practices were integrated into the work of care teams because capturing naturalistic behavior was important for measuring and monitoring a child's progress. For example, if the care team was helping a child manage his anger, they would review records over time to look for related changes in behavior, such as more instances of describing his emotions, more instances of asking for a break when frustrated, fewer uses of inappropriate language, or less time needed for the child to calm down during an emotional outburst.

The challenges of representing behavior in the record included applying standardization and quantification to subjective aspects of a person's health; and the need to observe naturalistic behavior, which required mobility and could lead to a significant amount of data transfer. Below, we outline key practices and artifacts of record keeping, and later we describe their role in the tensions we identified.

4.1. Standardization

The behaviors being monitored were highly individualized based on each child's unique goals. However, to allow for some standardization and consistency across care teams, four school-wide behavior categories were implemented: be safe, use kind words, complete work, and follow directions. These categories were helpful when used as rules and expectations, which were communicated frequently to all of the children as positive behavioral supports. However, when these same categories were applied for standardizing the behavioral record, they affected data qualities such as granularity and meaningfulness that are crucial for reflection among the care team (Marcu et al. 2019). Therefore, in practice, a care team would see that the child working on managing his anger would have changes in the 'be safe' or 'use kind words' categories, without always knowing specifically how his behavior changed or in which contexts. Care teams commonly discussed this type of data in order to elicit or infer such details from the record keeper and other team members most familiar with the child's behavior.

4.2. Quantification

As mentioned, behaviors were measured using frequency and duration. Frequency would be obtained using tally marks for a discrete instance of a behavior such as raising one's hand, hitting another student, or a defiant comment toward staff.

Duration would be captured by glancing at a clock on the wall, or sometimes using a timer, for prolonged instances such as tantrums or refusing to participate in an activity. Supervisors had created a single-sided data sheet on which classroom staff recorded behavioral data for all of their students for the full school day. This data sheet was carried primarily by the teacher and shared among the classroom staff on a clipboard, such as the red one in Figure 2b. A matrix design allowed for viewing the classroom's full roster students along the rows, and all periods or scheduled activities for the school day along the columns. Classroom staff recorded behaviors using tally marks in the box for the corresponding student and period. In addition, free-form annotations along the margins of this data sheet were sometimes used to note initials of which peer or staff a behavior was directed toward, or other contextual information.

4.3. Mobility

Placing the data sheet on a clipboard enabled the classroom staff to take it with them as they attended to each child around the classroom space, as well as outside of the classroom when activities required their students to be in other spaces such as the gymnasium, music room, cafeteria, and hallways. Activities taking place in other spaces were led by other staff in the school, while the teacher would often stay behind in the classroom to use the time to prepare lesson plans and activities. The paraprofessionals would therefore accompany the students and take the clipboard with them in order to maintain records on any behaviors occurring outside of the classroom. This type of ongoing record keeping was important because transitions between activities and spaces could trigger problem behaviors. However, teams were not always able to carry the clipboard with them, or share it between them at all times, so use of the data sheet was fluidly intermixed with other tools such as pieces of scrap paper and writing on one's hand, as shown

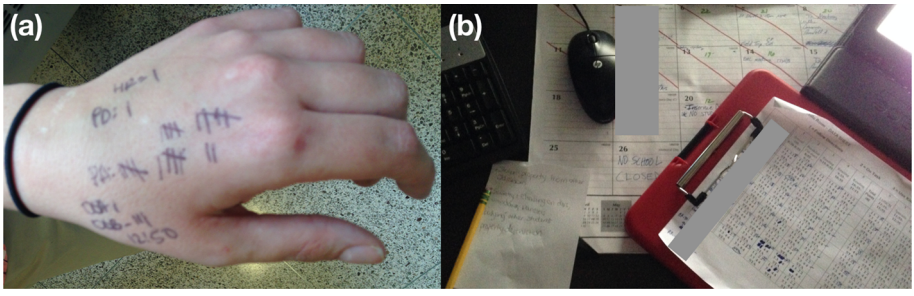

Figure 2. Care teams used a range of tools to record data about children's behaviors throughout the school day: **a** writing on their hand, **b** on scrap pieces of paper, and on a data sheet on a clipboard. The design process for our tablet-based Lilypad prototype sought to match the mobility of these tools and replicate the same structure for behavioral records, while reducing the amount of data transfer.

in Figure 2. The photograph in Figure 2b was taken of a teacher's desk with artifacts laid out in the middle of a school day. On the desk are a computer keyboard, a pencil on a piece of scrap paper, a computer mouse, a paper calendar, the school's standard data sheet on a clipboard, and a tablet computer provided to all employees by the school.

4.4. Data transfer

The clipboard data sheet served as the primary point of capturing behavioral data in the moment, and then was central to the transfer of data across several types of records. This process was complicated in part by the need to transfer and prepare data in different ways to meet the needs of various stakeholders. When a team member jotted down data on a piece of scrap paper or on their hand, they would transfer these data to the clipboard later that day once they had access to it. At the end of the day, the teacher, or primary record keeper, was responsible for transferring summarized data (e.g., frequency and duration totals, a few sentences describing highlights) to web-based case management software for the purposes of billing to insurance. Every few weeks, the primary record keeper would also spend several hours preparing reports for auditor visits or parent meetings, by transferring data from the clipboard to formats including a point graph on paper, spreadsheet, word processor, and web form.

5. Care team member roles with the record

Certain members of the care team carried the record keeping responsibilities, because this required direct observation of behavior. The rest of the care team was involved in some way with ongoing treatment decisions, for which they made use of the behavioral data in assessing the effects of interventions on behaviors. Behavioral goals were set to help monitor progress over time and also report to those outside of the organization, including parents and auditors. Interactions across care team member roles occurred through (or sometimes around) the record, and often tensions resulted from the ways records were created, shared, and used in care coordination.

5.1. Child and home caregivers: largely excluded from intraorganizational record

The children themselves, and their home caregivers, will naturally hold insights that are complementary to those of professional staff on the care team. However, we found that they are largely excluded from work with the record, because record keeping practices were a part of intraorganizational processes shared among its staff. Our findings reveal missed opportunities for incorporating behavioral data from children and home caregivers into the record. In one exception, some children, who had made significant progress and were consistently meeting

behavioral goals, were tasked with capturing separate data for self-assessment. Using a laminated card, they would rate their own behavior and then check-in with their teacher periodically to verify their self-assessed behavioral data. This practice was used as a way of promoting self-management and independence, and not intended to contribute to record keeping. The teacher could use this practice to inform their record keeping, but the child was not otherwise involved with the creation of their records. Similarly, parents and other caregivers in the home are observers of a child's daily behaviors. Although there was no mechanism for them to contribute to the working record, they would occasionally share information about a child's behavior via phone calls or in person, and that information could indirectly become a part of the record through another member of the care team.

5.2. Immediate care team: primary record keepers

Responsibilities for record keeping primarily rested with those members of the immediate care team who had the most opportunities to observe a child's behavior—the classroom staff, comprised of a teacher and several paraprofessional educators. Those who spent most time with the child had the most intimate understanding of their day-to-day life, personality, progress, and needs. They leveraged this understanding to interpret behavioral incidents and stay in tune with changes in behavior, so that they could record them. These individuals served as the instrument of measurement in recording data so they could monitor behaviors they were working to increase (e.g., following rules, emotional self-management, appropriate social interactions) or decrease (e.g., destruction of objects, self-harm, aggression toward peers). They would make quick decisions about how to respond to behaviors in the moment, doing their best to apply interventions and strategies agreed upon by the care team, while also maintaining the record with real-time capture of behavioral data.

5.3. Extended care team: users of record for corroboration

We found that even some professional members of the care team, who were part of organizational processes, still did not have the ability to contribute to the working record. Only classroom staff, who had daily contact with the child, held the responsibility for maintaining records used by the entire care team. This meant that others who had interactions with the child, and even recorded some behavioral data, did not contribute these data to the team's working record. For example, in-house speech therapists worked with a child in one-on-one therapy sessions once a week. The therapist picked up a child from the classroom during a scheduled period to work with them in their personal office. In these sessions, they recorded data for their own use. Then, during brief conversations with classroom staff as they were picking up or dropping off the child to the classroom, they might share thoughts about behaviors the had observed. However, their coordination with the

rest of the team occurred largely via reflecting on the working record maintained by the classroom staff.

Less proximal members of the extended care team had little to no direct contact with a child, and therefore were the most distant users of the behavioral data in the record. The record gave them a perspective into a child's behaviors, in the near or total absence of their own direct observation. These roles, including supervisors and psychiatrists, drew on significant training and a breadth of experience that enabled them to consult on a range of cases. They had the lowest proximity to data from individual children's cases, because they had the largest caseloads and tended to oversee programs, spending limited time directly observing each case. Instead, they reviewed data recorded by others in order to check on progress, and performed their own observation as needed in order to provide consultation. This made them dependent on other members of the care team to record accurate and useful data that they would be able to use for making treatment decisions. They also asked clarifying questions about data, including asking others for more contextual details to help them understand specific incidents or overall trends. For example, they would try to understand if there was a typical antecedent to a behavior that could reveal why that behavior may be occurring. Applying their training and experience with best practices, their role on the team was to interpret the data in order to provide diagnosis, prescriptions, treatment plans, and other high-level direction.

5.4. Distributed stakeholders: distant consumers of the record for oversight

Finally, more distant stakeholders could influence practices even if they were not part of the care team. Coordinating across organizational boundaries could require sharing of records, and health care policy could require certain data to be captured in records. For example, supervisors were responsible for ensuring adherence to data requirements from insurance companies, and passing these on to record keepers. The ways these distant requirements and needs were translated across the care team and impacted practices sometimes created tensions, which centered around the burden on record keepers. Certain therapists served as family liaisons for the care team, and in wanting to report accurate information to parents, they would depend on requesting the most updated records from teachers. Likewise, the day before an audit, which depended on thorough and up-to-date records, teachers had to take time away from their classroom in order to find a quiet place to sit at a computer and work on records. These occurrences highlighted the tension between their responsibilities to the children in their care, as well as maintaining records for the rest of the care team. Teachers served as record keepers and had to prepare reports for these different stakeholders, but they struggled with how much these activities took time away from their direct work with children.

6. Origins and design of the Lilypad prototype

Our formative fieldwork gave us an understanding of the challenges involved in representing behavioral data in the working record. A prototype computerized working record, Lilypad, was also a product of our formative fieldwork and action research in partnership with three care teams.

6.1. Motivations to adopt a novel electronic health record

Fairly early in our fieldwork, our partners within the school began drawing our attention to the challenges of record keeping—from the design of their data sheets, to the amount of data transfer required, to the usability problems with their reporting software. Those responsible for record keeping found it challenging to manage these practices with their concurrent care responsibilities in direct contact with multiple children. Those reviewing and consuming the records sometimes wanted more detailed and immediate access to data, with less dependency on others.

6.1.1. Need for support to achieve real-time data collection

Supervisors generally had the expectation that behaviors were recorded in the moment as they occurred. Maintaining records in real time would follow evidence-based practice, and paper-based practices had persisted in large part to enable the mobility and flexibility for always having the record on hand. However, recording accurate data in the moment was challenging because the record keepers were also those members of the care team most responsible for managing children's behaviors throughout the day and applying agreed upon interventions in response to those behaviors.

Record keepers adopted their own practices to help them update records as quickly as they could after observing a behavior. For example, a paraprofessional educator followed a child outside of the classroom during an incident, and did not have access to any paper for continuing to record the child's behaviors. She did have a pen, however, which she used to record several types of data on her hand as shown in Figure 2a: tally marks within behavioral categories (e.g., 'PD' for property destruction), and a notation of the time (e.g., '12:50') for monitoring the duration of the incident. In the absence of these types of workarounds, classroom staff would record behaviors later in the day when they had both access to the data sheet and a free moment to add to it. It was not uncommon for this opportunity to occur hours after a behavioral incident had taken place, requiring classroom staff to rely on memory in recording data on the incident.

6.1.2. Costs of manual data transfer

After collecting data on behaviors in the moment, record keepers were tasked with transferring the data multiple times to meet the needs of the care team, both intraorganizationally and interorganizationally. Supervisors set requirements

for reporting to insurance companies, including selecting the web-based case management software and passing along requirements for how data should be entered for billing. Insurance companies required specific information for billing purposes, such as detailed logs of activities that constituted behavioral intervention and treatment. Each member of the care team was required to account for their work with each child, sometimes down to the minute. Daily narratives were also needed for each child, outlining their treatment and progress, which was a time-consuming activity. Audits were common, sometimes once per year, adding pressure to create detailed and accurate records so as not to create problems for oneself or the organization. Yet record keepers often felt they did not have the time to complete these, admitting to copying and pasting re-used narratives when they knew this practice was not permitted.

To the record keepers, manual data transfers and calculations required a significant amount of time, causing them some concern about this time being taken away from their work with children. They spoke of attempting to maintain their focus on the children, which left no time for what they referred to as clerical work: 'there is no clerical time within the 40 hour work week, you do billing and paperwork on your own time'. However, regularly taking paperwork home resulted in burnout, causing some staff to instead allocate a dedicated amount of school time away from their students for managing data. Record keepers spoke about the logistical challenges of balancing all of their work, because the task of creating and maintaining records rested with the same staff members primarily responsible for ensuring children's safety and meeting their behavioral needs.

6.1.3. Dependency for access to data

Other members of the care team were dependent on the record keeper for access to records. Records were largely stored in binders on a bookshelf in the classroom, where they were not easily accessed or shared. Members of the care team would either have to wait for the next report to be prepared, or ask for the appropriate binder in order to review data. Figure 3 shows a monthly team meeting for which the teacher has prepared a report that each team member is reviewing on their own computer. The teacher, in the foreground, has also brought several binders of data which includes point graphs they have created by hand. These materials were only referenced during meetings by the teacher themself, as the others reviewed the summarized report during discussion of team members' interpretations regarding the child's progress. Therefore, care teams depended on the record keeper to manage and prepare the data, as well as summarize and describe the data when others did not have the time or ability to review records.

6.1.4. Existing technologies not fit for purpose or user friendly

Record keepers described past efforts to manage data using the tablet computers provided by their organization, which were currently used for email and viewing documents. These efforts were motivated by the desire to reduce time dedicated

Figure 3. A care team meeting with records in the form of point graphs filed in binders, and summarized reports reviewed digitally on individual computers. All identifiers have been removed.

to paperwork, but were ultimately unsuccessful. Existing tablet applications available for download had significant usability issues, and did not allow for the flexibility and customization required to capture children's behavioral incidents on a daily basis within a classroom setting. For example, some applications were designed for one-on-one behavioral therapy sessions, and some required so much navigation to customize the records for each child that the interaction design was infeasible for capturing data in real-time. Record keepers expressed ongoing frustration with the demands that paper-based records placed on them, and continued to look to their tablet computers for a possible solution.

6.2. Lilypad concept and interaction design

The concept for Lilypad evolved through continued fieldwork, and we elicited input from care teams throughout the design process. On multiple occasions during the design process, inconsistencies were uncovered in the implementation of practices with the paper record. At times during the design process, the tensions we had observed would manifest as contradictory preferences between supervisors, who tried to enforce ideal recording practices, and behavioral specialists working in the classroom, who recorded data as they could in the moment. Sometimes care teams were unaware of these contradictions, and other times we acted as mediator to find a compromise between what was ideal and what was more realistic.

Informed by our fieldwork, we focused on the challenges of entering behavioral data into the record. Record keepers had highlighted to us the difficulties they experienced with existing tools, and other members of the care team were interested in using the records they created. Over time we narrowed in on the work of

record keepers, and addressing tensions from their point of view. Ultimately, in order to manage the practical challenges of record keeping, we often prioritized the needs of record keepers in making design choices. As we will discuss later, this may have maintained the status quo, while addressing the creation and use of records across the full care team may have provided an opportunity to more holistically evaluate their coordination practices. For example, an alternative approach could have been understanding data quality and reporting workflows from more perspectives, and then aligning how we designed data collection more strategically toward those end goals. Instead, we viewed these aspects as a part of our long-term vision, and a future step in our design work.

The concept for a working record that we prototyped became known as Lilypad, to connote an interconnected ecosystem of devices that would provide each member of the care team a window into the data (as distinct and simple as a lily pad on the surface of water) while the system aggregated data from multiple sources (as with a network of intertwined roots underneath the water). Our aim was to manage the complexity of these connections to draw from a common database, so that each individual's interface would present a care team member with a usable record that fit the needs of their particular role. The Lilypad system is a network of tablet computers enabling care team members to record behavioral data with mobility, while also supporting the sharing and use of data as soon as they are entered into the record. Lilypad is designed to serve as a visualization of a personal history, similar to work such as LifeLines (Plaisant et al. 1996).

The interface for creating and maintaining records in real-time was designed to mimic the existing practice of carrying a clipboard with a data sheet for all of the children and all of the day's activities in one classroom (as shown in Figure 2b). Figure 4 shows the main Lilypad interface for recording data. Care teams view records by classroom—the top left corner of the screen indicates data are being recorded for the fictional Room 214. Of the list of students in that classroom shown in the navigation pane on the left side of the screen, the pseudonym Benedict is selected, and this child's record is shown on the rest of the screen. This interface was designed to enable quick switching between children's records. In a classroom setting, care teams are often working with multiple children at a given time, and are responsible for simultaneously creating records for all of their behaviors.

We observed that despite the flexibility of paper-based records, the standardization of how behaviors were recorded caused limitations in how informative the data were later for reflection. Our goal with Lilypad was therefore to create enough structure to enable collaborative reflection across care teams in the organization, while maintaining the possibility for behavioral categories to be easily customized for each child, and adjusted during the course of care. Three data types emerged as our focus for interaction design, because they were the most common practices for recording behavioral incidents: frequency, duration, and qualitative notes.

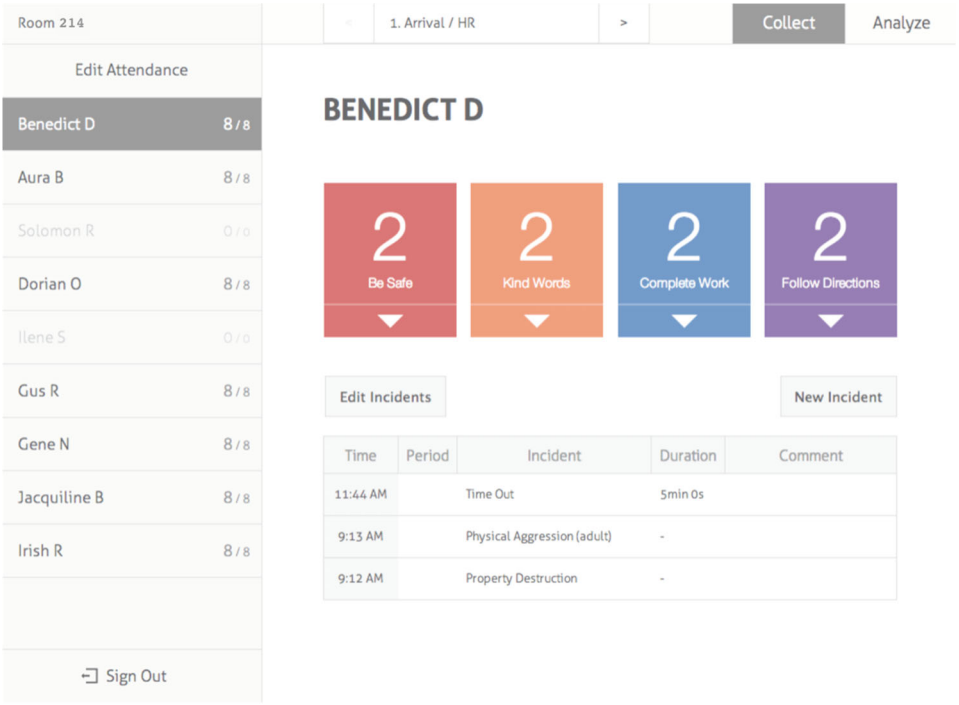

Figure 4. The Lilypad interface for creating records in real-time on a tablet, designed to support concurrent observation of behaviors across all children in a classroom. All identifiable information has been changed.

- *Frequency.* The top half of the record screen features counters to track the frequency of four to six types of behaviors. In Figure 4, Lilypad is customized with four counters for the school-wide behavioral rules.
- *Duration.* The bottom half of the screen features a log of duration-based behaviors such as time-outs. Users select an incident type from a custom list (e.g., physical aggression, out of assigned area), and enter a start time and end time. The times are used to calculate and display the duration of the behavior, and the period or activity during which the behavior occurred.
- *Qualitative notes.* Each time a behavior is recorded, the user can type a free-form note with qualitative details such as an antecedent event to the behavior, or names of other children or staff involved.

Focusing the record on these data types supported efforts to standardize practices using school-wide behavior categories, because these categories could be used as labels for counters that captured the frequency of each type of behavior. We also included additional frequency counters that could be customized to each child, and qualitative notes were meant to enable users to capture details that would help with interpreting a child's individual behaviors.

Once we determined together with the record keepers that the prototype was ready for use, we supported a transition and began a deployment study for 24 weeks, to continue understanding the challenges of capturing behavioral data.

7. Tensions in capturing and using behavioral data in an electronic health record

Insights from the deployment study helped us understand how the transition to use of Lilypad began to meet the needs of record keepers who had been looking for technological support—while also highlighting the tensions that persisted when it came to capturing behavioral data within records. We discuss how these tensions appeared before and after the deployment. We also reflect on the design decisions behind Lilypad and to what extent they affected how care teams used behavioral data in coordinating care.

7.1. Designing for existing practices supports adoption but can maintain status quo

In light of past challenges with utilizing their tablets for data collection, we were able to lower some barriers to adoption and integration of Lilypad by working together on its design so it could fit their workflow. Because we designed the Lilypad interface to closely mimic their existing record keeping practices, care team members were able to learn to use it quickly and while on the job. For each of their first three days of use, we conducted three to four hours of observation to be available for support during the transition, but they needed little to no guidance. By the second week of use, Olivia felt she was adjusting to Lilypad in her daily workflow as a record keeper, saying: 'it's nice that this [system] is becoming a part of my day'. A few months later, Olivia was preparing for maternity leave and a new care team member shadowed her for two weeks in preparation for filling her role. This new team member organically learned to use Lilypad as part of her training within three days, without technological support.

Additional team members, already accustomed to the organization's records, intuitively used Lilypad to enter a behavior they had observed into a child's record. This suggests the possibility of distributing responsibility for record keeping, as no training was required to use the system if the user was already familiar with the content of the record. Olivia also spoke of Lilypad as the type of support that she had been wanting her organization to provide: 'this is great, I've been asking for something [to help me with records] for a while, and they wouldn't even let me use Excel'. This sentiment speaks to the importance of support for those who carry the most responsibility for record keeping, and the lack of power they sometimes feel. Moreover, we saw that the content of the record, and the everyday workflows associated with it, could remain largely unchanged as we integrated a new tool for creating and using the records in real time.

By prioritizing a smooth transition to an electronic health record, to some extent we maintained the status quo. Our approach was to match existing practices as a starting point for adoption, focusing first on the needs of record keepers who carried the greatest responsibility and costs for data collection. However, once they had adopted Lilypad, there was not necessarily a clearer path to evaluating the effectiveness of data collection practices. Our design choices met the needs of record keepers, but did not always address tensions around navigating what kind of behavioral data makes for a useful record. For example, a conversation we observed during a meeting with their earlier use of paper-based records revealed an issue fundamental to their record system which we did not address:

> Psychiatrist: 'He seems to be doing better—based on this data, and the brief moments I've seen him in the classroom and hallway. He seems less irritable'.
>
> Record keeper: 'This is an example of how our record system doesn't work for someone like him, because it's showing something else ... it should be telling us, when he is supposed to be on task, is he?'
>
> Mental health therapist: 'The data may be accurate but it's not reflecting how he's really doing. I'm wondering if we can separate that behavior out and record that separately'.
>
> Record keeper: 'That would be hard'. [Laughs nervously]
>
> Mental health therapist: 'I'm just thinking, if an auditor came in here and looked at the numbers, it would not be reflected that he's really where he is ... how can we better reflect that he still needs to be in this program? Somebody from the outside may think, oh he's doing great, why is he still here?'

This discussion shows how members of the care team were concerned with demonstrating to those outside of their organization whether a child's needs were matched to their program and whether this placement (which may be more expensive for the school district than others) was appropriate. The record keeper's nervous response in this conversation was a common refrain to the occasional suggestion that more data would make a record more representative. Without a flexible way to adjust how records were created locally in order to more effectively represent behavioral needs to distant stakeholders, collecting more data was the only option. As we will discuss, Lilypad made it easier for record keepers to manage data they would bring to such meetings, potentially even facilitating more detailed records, but did not enable them to adjust standardization and reappropriate the record in a way that addressed this tension.

7.2. Data quality efforts can compete with intraorganizational coordination

Rather than having a uniform standard for data quality, there was tension between what level of quality was needed interorganizationally versus intraorganizationally.

The conversation in the previous section illustrates why care teams were concerned about producing high quality data for those outside of their organization. Intraorganizationally, however, they relied upon the data less. Daily data collection practices were not often scrutinized by care teams, because they relied on their own and each others' professional judgement. Efficient data collection practices were a priority in order to focus attention on delivering care and have more time with each child, so efforts to improve data quality could be in direct competition with this priority. Uses of data were also not always aligned across members of a care team, and so data quality was not necessarily discussed. As a result, routine intraorganizational interactions with the data could have been missed opportunities to address data quality collectively (rather than by record keepers alone). An electronic record could potentially generate such opportunities, but again Lilypad maintained the status quo.

As described earlier, standardization in record keeping was avoided because it would not enable individualized and adaptive services, but a trade-off was inconsistency in how behavioral data were captured. The primary method of standardizing practices was in setting common expectations for children's behavior, and structuring record keeping in relation to those expectations, by using four school-wide rules: be safe, use kind words, complete work, and follow directions. For instance, members of the care team would respond to violence toward a peer or destruction of objects by physically and verbally intervening, reminding the child how to be safe, and then recording the incident using the corresponding behavioral category of 'be safe'. However, there could be varying interpretations of an incident across members of the care team, as well as different ways of intervening, which would ultimately result in different data. The varied number of individualized behaviors monitored at a given time within a classroom meant there were rarely operationalizations of behaviors that could be referenced. Adding to the challenge of maintaining consistency across record keepers, the four behavioral categories included some overlap—for example, a student who was not completing his work was likely also not following directions.

The transition to Lilypad upheld the designated structure for data collection within these categories, but the breadth and vagueness of the categories themselves remained. Supervisors sometimes used training and inter-rater reliability checks to maintain data reliability, and an electronic health record could be used to support these strategies, but they did not emerge as a priority at this stage of our project. With improved access to data collected across the team in real-time, we had intended to enable more self-comparison among record keepers so that they could identify inconsistencies and work toward operationalizations. However, we found that their focus was primarily on creating records for use by those with less proximity to the classroom. Intraorganizationally, among team members who worked together closely, records were less important for communicating and coordinating with one another. Our findings match those of past work which indicated

early childhood interventionists prefer to rely on their own observations and memory rather than go through the effort of collecting data for their own understanding (Sandall et al. 2004). Moreover, if Lilypad relieved any of the effort having to do with record keeping, those users were much more interested in directing their focus to the children, rather than the children's data.

7.3. Unstructured data are useful intraorganizationally but captured inconsistently

Unstructured data in the form of ad hoc notes appeared to be useful intraorganizationally for reflection, but it was captured inconsistently by record keepers. Without the ability to standardize this aspect of the record, and without uniform standards for data quality, use of Lilypad did little to make these practices more consistent. However, some record keepers who were already using this practice may have increased the level of detail in their notes. Lilypad also made it easier for other members of the immediate care team to review these data in real-time, without data transfer. It remains unclear to what extent this real-time access can address the tension between the usefulness of unstructured data and the effort required to routinely and consistently capturing these data.

For intraorganizational meetings, unstructured data provided the contextual information that care team members would often seek out from the record keepers and others with more proximity to the child's behavior. A record of an incident could be as minimal as adding to a running tally of the number of times a behavior had occurred that day. But a record could also include details such as the timed duration of the behavior, names of peers involved, or the antecedent event right before the behavior occurred, which could provide insight into the psychological trigger of the behavior. Due to the subjective nature of the data and the challenge of capturing it while also managing behaviors, the amount of detail in a record varied across care team members, based on their workload and personal work style.

In a transition from paper and pencil to a tablet computer, we were concerned that the task of record keeping could be hindered by typing with an on-screen keyboard. However, record keepers were able to maintain a similar amount of detail. For example, Jonah was one of the most meticulous record keepers, and typed many free-form notes into Lilypad the same way he had written them by hand. Jonah tended to type a note with many entries. He noted specific details of a behavior, for example qualifying a 'Property Destruction' incident with 'ripped up notebook, broke pencils'. He would also note the antecedent to a behavior, for example an 'Elopement Out of Area' incident occurred when the child 'walked out of room after destroying property, swearing at staff.' On the other hand, it was common for Claire to be so occupied in her classroom that she was not able to record data in the moment. As she had done on paper, Claire was able to catch up on recording data while using Lilypad. Because it did not enforce a specific

workflow or structure, Lilypad allowed Claire to record data when she could. Jonah was given the flexibility to add as many details as he wanted without being constrained, while Claire had the flexibility to catch up on records as best she could. In focusing on record keepers, we designed Lilypad to support local needs at the point of capturing behavioral incidents.

The rest of the care team was interested in reviewing these details that were captured in the record. In one meeting, the care team was reviewing a child's data using the Lilypad interface projected onto a wall-mounted display via Apple TV. Looking at the structured data represented in line graphs did not produce much discussion. However, they then turned to the log of incidents which showed more detail about the child's behaviors, rather than high-level trends. These types of details had not been typically transferred when they were using paper-based records, and so had been unavailable to the rest of the team in the reports that drove discussions in meetings (recall the binder in Figure 3). Viewing the Lilypad screen in Figure 5, a mental health therapist noted that three 'Time Out' incidents over the course of several days were all annotated with 'FD,' indicating the time out was a consequence of the child not following directions they had been given. She asked the rest of the team about this pattern, the potential causal relation, and what intervention might be applied more preventatively when the behavior first occurs, so as to avoid escalation to the point of requiring a time out. Such scenarios illustrated how collocated care teams could use unstructured data for reflection, but they were dependent on individual record keepers taking the time to enter these notes in a useful and consistent way.

7.4. Real-time records can shift responsibility for reflection thereby requiring renegotiation of roles

Interestingly, the mental health therapist above was not involved in the child's case, so she did not have any knowledge of the treatment plan or behavioral goals. She did not have much proximity to this child or their data, yet she was able to contribute a useful question to the team's collaborative reflection on the data. This was one indication that use of Lilypad caused a shift away from the sole responsibility on the teacher, as primary record keeper, to present the data to the rest of the team. Instead, the data were reviewed collectively as a group with different team members providing interpretations. These types of interactions show the potential of shared data representation to engage more members of the care team in reflecting on a child's progress. We viewed this as a positive change in the transition to Lilypad, however we reflect on how such a change can require renegotiation of roles across a care team, which may be disruptive or time consuming.

Across team member roles, there was perceived value in having graphs available in real-time as data were collected. For example, Claire was the record keeper most resistant to adoption due to previous experiences, yet she became noticeably more motivated after seeing the functioning graphs. She saw a concrete benefit

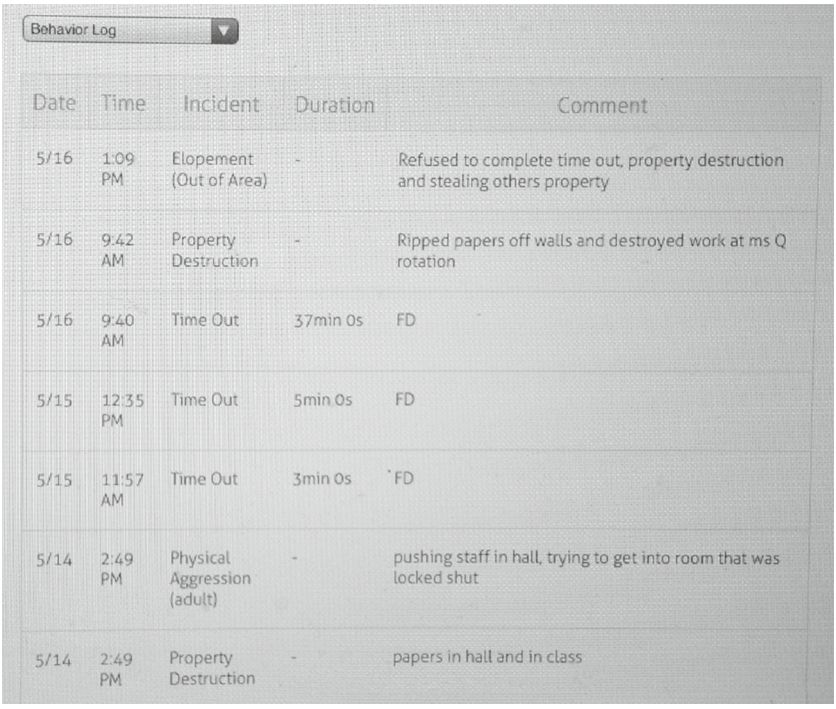

Figure 5. A sample behavior log in Lilypad displayed in reverse chronological order, with duration of the incident where applicable, and any free-form notes that may have been included in the entry to provide detail of the behavior(s) involved in the incident.

in not having to generate those graphs herself. A new mental health therapist was hired partway through the deployment, and she had a positive reaction when introduced to Lilypad's graphs. She requested to become a user because of the ability to review the data with others as an aid to communication about a child's behavioral progress. Other team members were similarly positive about the graphs helping with challenges of reporting to stakeholders outside of the organization, stating for example that the graphs would be easier for parents to understand than existing reports. Effective coordination with parents is important for a child's behavioral outcomes, however greater access to data has the potential to increase collaborative tensions, especially if the data are primarily intended for intraorganizational use (Marcu et al. 2019).

The discussions we observed around Lilypad data we hopeful that it could reduce dependency of other team members on record keepers for understanding a child's behaviors and progress. But the interest in improving access to data through real-time records has implications for how the care team works together, especially with members who are more distributed and coordinate interorganizationally. Once a member of a care team has easier access to a child's record, expectations for their review of the data and their involvement in the child's

case may change. Therefore, the role of care team members would need to be renegotiated as access to records changes.

7.5. Shared representation of data versus multiple representations of data

The transition to Lilypad looked toward improving behavioral monitoring over the long-term. This need was reflected, for example, in one record keeper's struggle to look at enough data within a binder at once during a meeting, at one point folding one piece of paper over a second in order to see more rows simultaneously of an extended table. Identifying an appropriate representation of the data, however, was a challenge.

Given the value placed on monitoring by multiple team member roles, we began some ideation with a broader range of data representations. During the process of envisioning new graphs and charts, team members found it difficult to think of the data in new ways. For instance, we suggested visualizations supporting identification of trends in the data such as correlations, but team members did not want such features because they did not match how they were used to reviewing their data. This resistance shows that care teams develop strong mental models of behavioral data, and there will be a barrier to adoption if new systems do not represent data in ways that directly map onto these mental models.

We found that even small changes in how data were represented made a significant difference to team members. Care teams discussed quantitative behavioral data in one of two ways: by absolute points earned (20/40) or percentage of points earned (50%). Different team members were discussing data based on their personal preference for one of the two representations. Over time, as we debated which representation we should use in Lilypad, they realized that each of the two representations revealed different aspects of a child's progress. Percentages helped them review data long-term because percentages accounted for variations in points possible (for example, due to the school scheduling a half-day or a student arriving late). Absolute points were harder to reflect on long-term because the denominator could change, but absolute points gave a more complete picture of a particular day or short-term period.

We might conclude, therefore, that multiple representations of the data can serve different purposes, whether supporting long-term and short-term practices, or adhering to the needs of individual team members. An alternative approach would be to focus design efforts on aligning team members' understanding of the data, from operationalization through data collection, to review of visualizations. In this way, shared representations could serve as common ground for reflection and coordination. Others have argued that team members should have a tailored view of shared information based on their personal involvement and context of use, though these studies took place within medical care (Amir et al. 2015; Bardram and Bossen 2005). Our findings suggest that representations of behavioral data may not have as clear a role for coordination, that care teams may prefer

to rely on professional judgement, and there may not be standard mental models used for reflecting on the data.

8. Discussion

Taking an action research approach, we examined how behavior is represented in the health record. We focused on the challenges of record keeping during observation of behavior, leading to the design and deployment of our Lilypad prototype. Investigating the transition from a paper-based record to Lilypad, and studying technology in use, we found persistent tensions surrounding representation and use of data in the record. We identified tensions related to the need to structure work with and around the record, while maintaining the flexibility and unstructured aspects of work that are key in behavioral health.

Bowers et al. discuss structured and unstructured work as interrelated, underscoring 'the multiple considerations which impinge upon the acceptability of technology in actual contexts–considerations which often require very difficult trade-offs' (Bowers et al. 1995). They describe these trade-offs based on how technologies help to organize workflow *from within* the organization and workflow *from without* the organization. Although we did not examine workflow, our study similarly explored intraorganizational and interorganizational work that has both structured and unstructured aspects. The record keeping practices we observed with behavioral data are largely not able to be standardized or sequential, and therefore do not fit definitions of workflow (Bowers et al. 1995; Dourish 2001). But we find that Bowers et al.'s discussion of tradeoffs aligns with the tensions that we identified. We therefore adapt their intra- and interorganizational distinction in order to explain tensions between the creation of records *from within* and the creation of records *from without*.

Records *from within* involved care team members reflecting on a child's behaviors individually and collectively in order to make adjustments to care based on the child's progress on behavioral goals. Often, however, these adjustments could happen through anecdotal reports based on trust in other team members' observations and professional judgement. As a result, creating records *from within* was less focused on capturing useful behavioral data, even though the records were generally valued and many processes revolved around the records. On the other hand, records *from without* were driven by accountability to stakeholders outside of the organization, related to billing, insurance, auditing, as well as placement decisions that balanced financial considerations with the individual needs of each child. Records *from without* were therefore seen by care teams to be more consequential, not only for financial accounting of work, but also because external members of the care team were much less familiar with the work of the organization and the behaviors and needs of each child.

For the purposes of reflecting on a child's progress intraorganizationally in the context of their program, care teams preferred leveraging the knowledge and

expertise of those members of the team with most proximity to the child. Other work has shown that early childhood interventionists don't feel they need to collect behavioral data for their own understanding, because they know the child well (Sandall et al. 2004). Our study shows this trust of observation, memory, and professional judgement extends to others on one's care team. That is, care team members may prefer secondhand accounts and interpretation of a child's behavioral progress over data. Meanwhile, even though care teams are concerned with interorganizational perceptions of what goes on in their program, creating records *from without* is so different to them from creating records *from within*, that the two become difficult to reconcile. Records *from within* tended to be created and used used more organically as part of reflection and other unstructured team work. Records *from without* were structured by reporting requirements set externally, and standardized practices and tools set internally by administrators and supervisors.

Arguably, designing and deploying the Lilypad system together could have been a potential opportunity to reconcile records *from within* and records *from without*. However, we prioritized ease of use and adoption for those team members most central to record keeping practices, and did not actively work to redesign the practices themselves. We knew that the organization had made previous attempts to transition away from paper-based records, unsuccessfully, and some on their staff were wary of another attempt. Focusing on feasibility, we also knew that existing reporting software was dictated *from without*, for example through a client management system for billing to insurance. The long-term vision was thus for Lilypad to eventually become interoperable with such systems, meeting local needs for record keeping *from within* first, and later facilitating workflows for reporting interorganizationally.

With increasing digitization of healthcare work, the difficulty of incorporating novel records such as Lilypad or even a basic EHR (Kaziunas et al. 2019) in behavioral health can be explained by the challenges of representing behavioral data, 'since data work often involves interpretation and assessments based on sense-making' (Bossen et al. 2019). There is a significant need to rely on a multidisciplinary care team for interpreting an individual's behavior and negotiating action with and around data in the record. We contrast the nature of this coordination with medicine, which uses data in the record at a larger scale for measuring organizational performance or even comparing services nationally (Pine and Mazmanian 2015). Behavioral data, on the other hand, is more suited to measuring individual treatment progress, and even representing that progress effectively can be difficult.

Medicine also draws from a single (or main) community of practice in determining how data will be used for accountability and decision-making (Pine and Mazmanian 2015). With regard to accountability, behavioral data are used to subjectively describe the nature of an individual's behavior so as to justify the need for treatment, given that objective measures of physical or biological symptoms

are often not possible. Accountability has more to do with documenting a need and rationale for why services are billed, as opposed to performance measures or protocols, which are difficult to define. Even when a behavior can be quantified, this requires a process of operationalizing the behavior of interest in a way that is difficult to standardize. Moreover, as multiple disciplines collaborate on a care team, each of them brings a different set of standards and each comes from a community of practice that may record and use behavioral data differently. The role of the record in coordination is therefore to enable ongoing interpretation and negotiation with one individual's behavioral data across communities of practice.

Dourish suggests that workflow technologies could serve to visualize work, whereby 'the role of the system is to find and present the order in the work, rather than, as traditionally conceived, to prespecify and enforce it' (Dourish 2001). Standardization and structure are key challenges with representing behavioral data, which is why natural language processing is one approach that has been applied with unstructured clinical notes in electronic health records, though it still carries significant limitations (Townsend 2013; Perera et al. 2013; Xu et al. 2007). To find and present order in the work through visualization, we could learn from how Lilypad was appropriated for these purposes. For example, when a care team used their Apple TV to project the Lilypad interface on a wall-mounted display in the meeting room, even without a dedicated visualization, a mental health therapist with little knowledge of the child's case was able to review recent behavioral data and reflect on what it might mean about the child's progress. This required a simple chronological log of the child's behaviors, pairing frequency and duration data with qualitative notes. Rather than sophisticated analytics or support for managing a large amount of data, Lilypad users wanted real-time access to review details of a child's behaviors in order to do their work. It may be that adjusting to an electronic health record over time would lead to use of more complex visualizations, but meeting these more straightforward data needs first enabled adoption and supported the collaborative work of the care teams.

9. Conclusion

In this three-phase study, we took an action research approach to understanding challenges with paper-based record keeping practices with behavioral data, then designing a novel electronic health record, and finally studying its use in a 24-week deployment. Our Lilypad prototype was designed to support the capture of behavioral incidents with three forms of data: frequency, duration, and qualitative notes. We have described how care teams struggle to distribute responsibility for managing behavioral data in the working record, with some members of the team more burdened and some more dependent. We identified five persistent tensions in representing behavioral data, which reveal the difficulty of manually capturing behavior in the record due to not only the nature of data that cannot be

standardized, but also the ways that intraorganizationl use of data can conflict with interorganizational use of data. Our deployment succeeded in facilitating transition from a paper-based to electronic health record, but in many ways did not change these tensions or always help care teams use data differently. Addressing these tensions requires understanding how data are captured and used from various perspectives across the care team, and rearranging workflows with data to provide more opportunities for interpretation and negotiation.

Acknowledgements

We thank our partners in this research for their time, expertise, and insights. We also thank reviewers for many helpful recommendations as we prepared this work for publication. This material is based upon work supported by the National Science Foundation under Grant No. IIS 1816319, Grant No. IIS 1029549, and a Graduate Research Fellowship to the first author.

References

Amir, Ofra; Barbara J. Grosz; Krzysztof Z. Gajos; Sonja M. Swenson; and Lee M. Sanders (2015). From Care Plans to Care Coordination: Opportunities for Computer Support of Teamwork in Complex Healthcare. In: *CHI '15. Proceedings of the 2015 ACM Conference on Human Factors in Computing Systems*, *Seoul, Republic of Korea, 2015*. New York, NY, USA: ACM Press, pp. 1419–1428.

Barbarin, Andrea; Tiffany C. Veinot; and Predrag Klasnja (2015). Taking our Time: Chronic Illness and Time-Based Objects in Families. In CSCW 2015. *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing, Vancouver, Canada, 14-18 March 2015*. New York: ACM Press, pp. 288–301.

Bardram, Jakob E.; and Claus Bossen (2005). A Web of Coordinative Artifacts: Collaborative Work at a Hospital Ward. In: *GROUP '05. Proceedings of the 2005 International ACM SIGGROUP*

Conference on Supporting Group Work, Sanibel Island, Florida, USA, 2005. New York: ACM Press, pp. 168–176.

Berg, Marc (2004). *Health Information Management*. Abington: Routledge.

Bonsignore, Elizabeth; Alexander J. Quinn; Allison Druin; and Benjamin B. Bederson (2013). Sharing Stories "in the Wild": A Mobile Storytelling Case Study Using StoryKit. *ACM Transactions on Computer-Human Interaction (TOCHI) - Special Issue of "The Turn to The Wild"*, vol. 20, no. 3, pp. 18:1–18:38.

Bossen, Claus; Kathleen H. Pine; Federico Cabitza; Gunnar Ellingsen; and Enrico Maria Piras (2019). Data work in healthcare: An Introduction. An Introduction. *Health Informatics Journal*. September 2019, pp. 465–474.

Bowers, John; Graham Button; and Wes Sharrock (1995). Workflow From Within and Without: Technology and Cooperative Work on the Print Industry Shopfloor. In: *ECSCW '95. Proceedings of the 1995 European Conference on Computer-Supported Cooperative Work, Stockholm, Sweden, 1995*. Dordrecht: Kluwer, pp. 51–66.

Charmaz, Kathy (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. London, UK: Sage.

Christensen, Lars Rune (2016). On intertext in chemotherapy: an ethnography of text in medical practice. *Computer Supported Cooperative Work (CSCW)*, vol. 25, no. 1, pp. 1–38.

Clarke, Alison J. (2011). *Design anthropology: Object culture in the 21st century*. Berlin: Springer.

Council on Children with Disabilities (1999). Care coordination: integrating health and related systems of care for children with special health care needs. *Pediatrics*, vol. 104, no. 4, pp. 978–981.

Dourish, Paul (2001). Process Descriptions As Organisational Accounting Devices: The Dual Use of Workflow Technologies. In: *GROUP '01. Proceedings of the 2001 International ACM SIGGROUP Conference on Supporting Group Work, Boulder, Colorado, USA, 2001*. New York: ACM Press, pp. 52–60.

Ellingsen, Gunnar; and Eric Monteiro (2003). Mechanisms for producing a working knowledge: Enacting, orchestrating and organizing. *Information and organization*, vol. 13, no. 3, pp. 203–229.

Erikson, Erik H. (1958). The nature of clinical evidence. *Daedalus*, vol. 87, no. 4, pp. 65–87.

Fitzpatrick, Geraldine (2004). Integrated care and the working record. *Health Informatics Journal*, vol. 10, no. 4, pp. 291–302.

Fitzpatrick, Geraldine; and Gunnar Ellingsen (2013). A review of 25 years of CSCW research in healthcare: contributions, challenges and future agendas. *Computer Supported Cooperative Work (CSCW)*, vol. 22, no. 4-6, pp. 609–665.

Greenwood, Charles R; Thomas R. Kratochwill; and Melissa Clements (2008). *Schoolwide prevention models: Lessons learned in elementary schools*. New York: Guilford Press.

Gunn, Wendy; Ton Otto; and Rachel Charlotte Smith (2013). *Design anthropology: theory and practice*. Abington: Routledge.

Hayes, Gillian R. (2011). The relationship of action research to human-computer interaction. *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 18, no. 3, pp. 15.

Hayes, Gillian R.; and Gregory D. Abowd (2006). Tensions in Designing Capture Technologies for an Evidence-based Care Community. In: *CHI '06. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Montréal, Québec, Canada, 2006*. New York, NY, USA: ACM Press, pp. 937–946.

Hayes, Gillian R.; Julie A. Kientz; Khai N. Truong; David R. White; Gregory D. Abowd; and Trevor Pering (2004). Designing Capture Applications to Support the Education of Children with Autism. In: *Ubicomp '04. Proceedings of the 2004 International Conference on Ubiquitous Computing, Nottingham, UK, 2004*. Berlin, Heidelberg: Springer, pp. 161–178.

Hayes, Gillian R.; Sen Hirano; Gabriela Marcu; Mohamad Monibi; David H. Nguyen; and Michael Yeganyan (2010). Interactive visual supports for children with autism. *Personal and ubiquitous computing*, vol. 14, no. 7, pp. 663–680.

Jagannath, Swathi; Aleksandra Sarcevic; and Andrea Forte (2018). "We Are Not Entirely Replacing Paper": Understanding Paper Persistence in Emergency Medical Settings. In: *CSCW '18. Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing*, *Jersey City, NJ, USA, 2018*. New York, NY, USA: ACM Press, pp. 249–252.

Johnson, Rose; Yvonne Rogers; Janet van der Linden; and Nadia Bianchi-Berthouze (2012). Being in the Thick of In-the-wild Studies: The Challenges and Insights of Researcher Participation. In: *CHI '12. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, *Austin, Texas, USA, 2012*. New York: ACM Press, pp. 1135–1144.

Kane, Bridget; and Saturnine Luz (2011). Information sharing at multidisciplinary medical team meetings. *Group Decision and Negotiation*, vol. 20, no. 4, pp. 437–464.

Kaziunas, Elizabeth; Michael S. Klinkman; and Mark S. Ackerman (2019). Precarious Interventions: Designing for Ecologies of Care. *Proceedings of the ACM on Human-Computer Interaction*, vol. 3, no. CSCW, pp. 113.

Malone, Thomas W.; and Kevin Crowston (1994). The Interdisciplinary Study of Coordination. *ACM Comput. Surv.*, vol. 26, no. 1, pp. 87–119.

Marcu, Gabriela; Allison Spiller; James E. Connell; and Laura R. Pina (2019). Breakdowns in home-school collaboration for behavioral intervention. In: *CHI'19. Proceedings of the 2019 ACM Conference on Human Factors in Computing Systems, Glasgow, Scotland, UK, 2019*. New York: ACM Press, p. 661.

Marcu, Gabriela; Kevin Tassini; Quintin Carlson; Jillian Goodwyn; Gabrielle Rivkin; Kevin J. Schaefer; Anind K. Dey; and Sara Kiesler (2013). Why Do They Still Use Paper? Understanding Data Collection and Use in Autism Education. In: *CHI '13. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, *Paris, France, 2013*. New York, NY, USA: ACM Press, p. 3177–3186.

Marcu, Gabriela; Anind K. Dey; and Sara Kiesler (2014). Designing for Collaborative Reflection. In: *PervasiveHealth '14. Proceedings of the 2014 International Conference on Pervasive Computing Technologies for Healthcare*, *Oldenburg, Germany, 2014*. Brussels: ICST, pp. 9–16.

Marcu, Gabriela; Anind K. Dey; Sara Kiesler; and Madhu Reddy (2016). Time to Reflect: Supporting Health Services over Time by Focusing on Collaborative Reflection. In: *CSCW '16. Proceedings of the 2016 ACM Conference on Computer-Supported Cooperative Work & Social Computing*, *San Francisco, California, USA, 2016*. New York: ACM Press, pp. 954–964.

Marcu, Gabriela; Allison Spiller; Jonathan Arevalo Garay; James E. Connell; and Laura R. Pina (2019). Breakdowns in home-school collaboration for behavioral intervention. In: *CHI '19.: Proceedings of the 2019 ACM Conference on Human Factors in Computing Systems, Glasgow, Scotland, UK, 2019*. New York: ACM Press, pp. 661.

Marshall, Paul; Richard Morris; Yvonne Rogers; Stefan Kreitmayer; and Matt Davies (2011). Rethinking 'Multi-user': An In-the-wild Study of How Groups Approach a Walk-up-and-use Tabletop Interface. In: *CHI '11. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, *Vancouver, BC, Canada, 2011*. New York: ACM Press, pp. 3033–3042.

Mentis, Helena M. (2010). *Emotion Awareness and Invisibility in an Emergency Room: A Socio-Technical Dilemma*. Ph.D. thesis, The Pennsylvania State University, University Park, PA.

Messeter, Jörn; and Daryn Molenaar (2012). Evaluating Ambient Displays in the Wild: Highlighting Social Aspects of Use in Public Settings. In: *DIS '12. Proceedings of the Designing Interactive Systems Conference*, *Newcastle Upon Tyne, United Kingdom, 2012*. New York: ACM Press, pp. 478–481.

Mønsted, Troels; Madhu C. Reddy; and Jørgen P. Bansler (2011). In: *ECSCW 2011. Proceedings of the 2011 European Conference on Computer Supported Cooperative Work*, *Aarhus, Denmark, 2011*. Berlin: Springer, pp. 81–100.

Østerlund, Carsten S. (2008). Documents in place: demarcating places for collaboration in healthcare settings. *Computer Supported Cooperative Work (CSCW)*, vol. 17, no. 2-3, pp. 195–225.

Park, Sun Young; Katie Pine; and Yulan Chen (2013). Local-universality: Designing EMR to Support Localized Informal Documentation Practices. In: *CSCW '13. Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, *San Antonio, Texas, USA, 2013*. New York: ACM Press, pp. 55–66.

Perera, Sujan; Amit Sheth; Krishnaprasad Thirunarayan; Suhas Nair; and Neil Shah (2013). Challenges in understanding clinical notes: Why NLP engines fall short and where background knowledge can help. *DARE '13: Proceedings of the 2013 international workshop on Data management & analytics for healthcare, San Francisco, CA, USA, 2013*. New York: ACM Press, pp. 21–26.

Pine, Kathleen; and Melissa Mazmanian (2015). Emerging insights on building infrastructure for data-driven transparency and accountability of organizations. In: *iConference'15. Proceedings of the iConference, Newport Beach, CA, USA, 2015*. Grandville, MI: iSchools, pp. 1–13. http://hdl.handle.net/2142/73454.

Plaisant, Catherine; Brett Milash; Anne Rose; Seth Widoff; and Ben Shneiderman (1996). LifeLines: Visualizing Personal Histories. In: *CHI '96. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, *Vancouver, British Columbia, Canada, 1996*. New York: ACM Press, pp. 221–227.

Saario, Sirpa; Christopher Hall; and Sue Peckover (2012). Inter-professional electronic documents and child health: A study of persisting non-electronic communication in the use of electronic documents. *Social Science & Medicine*, vol. 75, no. 12, pp. 2207–2214.

Sandall, Susan R.; Ilene S. Schwartz; and Betsy Lacroix (2004). Interventionists' perspectives about data collection in integrated early childhood classrooms. *Journal of Early Intervention*, vol. 26, no. 3, pp. 161–174.

Siek, Katie A.; Gillian R. Hayes; Mark W. Newman; and John C. Tang (2014). Field deployments: Knowing from using in context. In J. Olson and W. Kellogg (eds) *Ways of Knowing in HCI*. Berlin: Springer, pp. 119–142.

Thompson, James D. (1967). *Organizations in action: Social science bases of administrative theory*. Abington: Routledge.

Timmermans, Stefan; and Marc Berg (1997). Standardization in action: achieving local universality through medical protocols. *Social studies of science*, vol. 27, no. 2, pp. 273–305.

Townsend, Hilary (2013). Natural language processing and clinical outcomes: the promise and progress of NLP for improved care. *Journal of AHIMA*, vol. 84, no. 2, pp. 44–45.

Xu, Hua; Peter D. Stetson; and Carol Friedman (2007). A study of abbreviations in clinical notes. In: *AMIA'07 Proceedings of the AMIA Annual Symposium*, Chicago, IL, USA, 2007. p. 821–825.

Yell, Mitchell; Carl Smith; Antonis Katsiyannis; and Mickey Losinski (2018). Mental Health Services, Free Appropriate Public Education, and Students With Disabilities: Legal Considerations in Identifying, Evaluating, and Providing Services. *Journal of Positive Behavior Interventions*, vol. 20, no. 2, pp. 67–77.

Publisher's Note