Behavioral Monitoring to Identify Self-Injurious Behavior among Children with Autism Spectrum Disorder

Kristine Dianne Cantin-Garside

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Industrial and Systems Engineering

Maury A. Nussbaum, Chair
Sunwook Kim
Zhenyu Kong
Divya Srinivasan
Susan W. White

February 18, 2019
Blacksburg, Virginia

Keywords: Autism spectrum disorder, self-injurious behavior, accelerometers, classification, machine learning

Behavioral Monitoring to Identify Self-Injurious Behavior among Children with Autism Spectrum Disorder

Kristine Dianne Cantin-Garside

ABSTRACT

Self-injurious behavior (SIB) is one of the most dangerous behavioral responses among individuals with autism spectrum disorder (ASD), often leading to injury and hospitalization. There is an ongoing need to measure the triggers of SIB to inform management and prevention. These triggers are determined traditionally through clinical observations of the child with SIB, often involving a functional assessment (FA), which is methodologically documenting responses to stimuli (e.g., environmental or social) and recording episodes of SIB. While FA has been a "gold standard" for many years, it is costly, tedious, and often artificial (e.g., in controlled environments). If performed in a naturalistic environment, such as the school or home, caregivers are responsible for tracking behaviors. FA in naturalistic environments relies on caregiver and patient compliance, such as responding to prompts or recalling past events.

Recent technological developments paired with classification methods may help decrease the required tracking efforts and support management plans. However, the needs of caregivers and individuals with ASD and SIB should be considered before integrating technology into daily routines, particularly to encourage technology acceptance and adoption. To address this, the perspectives of SIB management and technology were first collected to support future technology design considerations (Chapter 2). Accelerometers were then selected as a specific technology, based on caregiver preferences and reported preferences of individuals with ASD, and were used to collect movement data for classification (Chapter 3). Machine learning algorithms with featureless data were explored, resulting in individual-level models that demonstrated high accuracy (up to 99%) in detecting and classifying SIB.

Group-level classifiers could provide more generalizable models for efficient SIB monitoring, though the highly variable nature of both ASD and SIB can preclude accurate detection. A multi-level regression model (MLR) was implemented to consider such individual variability (Chapter 4). Both linear and nonlinear measures of motor variability were assessed as potential predictors in the model. Diverse classification methods were used (as in Chapter 3), and MLR outperformed other group level classifiers (accuracy ~75%).

Findings from this research provide groundwork for a future smart SIB monitoring system. There are clear implications for such monitoring methods in prevention and treatment, though additional research is required to expand the developed models. Such models can contribute to the goal of alerting caregivers and children before SIB occurs, and teaching children to perform another behavior when alerted.

Behavioral Monitoring to Identify Self-Injurious Behavior among Children with Autism Spectrum Disorder

Kristine Dianne Cantin-Garside

GENERAL AUDIENCE ABSTRACT

Autism spectrum disorder (ASD) is a prevalent developmental disorder that adversely affects communication, social skills, and behavioral responses. Roughly half of individuals diagnosed with ASD show self-injurious behavior (SIB), including self-hitting or head banging), which can lead to injury and hospitalization. Clinicians or trained caregivers traditionally observe and record events before/after SIB to determine possible causes ("triggers") of this behavior. Clinicians can then develop management plans to redirect, replace, or extinguish SIB at the first sign of a known trigger. Tracking SIB in this way, though, requires substantial experience, time, and effort from caregivers. Observations may suffer from subjectivity and inconsistency if tracked across caregivers, or may not generalize to different contexts if SIB is only tracked in the home or school.

Recent technological innovations, though, could objectively and continuously monitor SIB to address the described limitations of traditional tracking methods. Yet, "smart" SIB tracking will not be adopted into management plans unless first accepted by potential users. Before a monitoring system is developed, caregiver needs related to SIB, management, and technology should be evaluated. Thus, as an initial step towards developing an accepted SIB monitoring system, caregiver perspectives of SIB management and technology were collected here to support future technology design considerations (Chapter 2). Sensors capable of collecting the acceleration of movement (accelerometers) were then selected as a specific technology, based on the reported preferences of caregivers and individuals with ASD, and were used to capture SIB movements from individuals with ASD (Chapter 3). These movements were automatically classified as "SIB" or "non-SIB" events using machine learning algorithms. When separately applying these methods to each individual, up to 99% accuracy in detecting and classifying SIB was achieved.

Classifiers that predict SIB for diverse individuals could provide more generalizable and efficient methods for SIB monitoring. ASD and SIB presentations, however, range across individuals, which impose challenges for SIB detection. A multi-level regression model (MLR) was implemented to consider individual differences, such as those that may occur from diagnosis or behavior (Chapter 4). Model inputs included measures capturing changes of movement over time, and these were found to enhance SIB identification. Diverse classification models were also developed (as in Chapter 3), though MLR outperformed these (yielding accuracy of ~75%).

Findings from this research provide groundwork for a smart SIB monitoring system. There are clear implications for monitoring methods in prevention, though additional research is required to expand the developed models. Such models can contribute to the goal of alerting caregivers and children before SIB occurs, and teaching children to perform another behavior when alerted.

Acknowledgements

Thank you so much to my advisor, Dr. Maury Nussbaum for all of your support, guidance and encouragement. Your advice and wisdom are invaluable, and you have greatly influenced me as a researcher and professional.

Thank you so much to my committee members, Dr. Sunwook Kim, Dr. Zhenyu Kong, Dr. Divya Srinivasan, and Dr. Susan White for your continued help, direction and mentorship.

Thank you to Dr. Shyam Ranganathan and Dr. Rupa Valdez for all of your time and support.

Thank you to Ligia Antezana for all of your time in the clinic and your expertise.

Thank you to Louie Buchen, Ashley Deer, Chung Do Kim, and Diogo Fortes for your help and interest in this research.

Thank you so much to my incredible husband, Jordan, for your continuous love, support and patience during this entire process. I could not have done this without you!

Thank you to my wonderful family for your love and support, especially parents (Michael and Susan Cantin), sister Alyssa (Dr. Cantin-Giles), and my grandparents ("if you think you will, if you think you won't, you're absolutely right"). I am so grateful for you.

Thank you to my peers and friends for all of your ongoing support (labmates, especially Dr. Helen Nogueira for her friendship from VT to Brazil, officemates, and GCF Women's Bible Study).

Most of all, thank You God for all of Your many blessings. I owe everything to You!

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Autism spectrum disorder (ASD) is a developmental disorder characterized by communicative, social, and behavioral impairments (APA, 2013), affecting in 1 in 59 youth in the United States (Baio et al., 2018). There is a growing need for access to interventions and long-term care for the described challenges, specifically in response to the nearly threefold increase in the estimated prevalence of ASD from 2000 to 2014 (Baio et al., 2018; Dawson et al., 2010; Lucyshyn et al., 2007; Rice, 2007). Care is particularly critical for self-injurious behavior (SIB), which affects ~50% of people with ASD and often leads to severe physical damage and hospitalization (McTiernan, Leader, Healy, & Mannion, 2011; Minshawi et al., 2014; Murphy, Healy, & Leader, 2009; Richards, Oliver, Nelson, & Moss, 2012; Rooker et al., 2018). With diverse behavioral presentations, definitions of SIB can vary (Lam & Aman, 2007; Rojahn, Matson, Lott, Esbensen, & Smalls, 2001). The following definition of SIB reflects prior work (Iwata et al., 1994; Pace, Iwata, Edwards, & McCosh, 1986; Rojahn et al., 2001), and is used for the purpose of this research:

> SIB is an action that could cause harm or pain to the self and may serve different functions. SIB excludes accidental incidents (e.g., tripping and falling) and behaviors intended for cultural or aesthetic reasons (e.g., ear piercings). SIB itself may or may not be performed with the intention to harm. It can be repetitive, and can include classic behaviors (e.g., head banging, self-hitting or biting), fine behaviors (e.g., picking or pinching), and potentially injurious body contact with objects or the environment (e.g., repeatedly hitting into walls). An "episode" of SIB is defined as a continuous time period of a specific behavior type, of SIB or of co-occurring SIB, preceded and followed by other types of SIB or other non-SIB actions.

SIB may have a range of functions, which can differ within and between individuals (Iwata et al., 1994; Minshawi, Hurwitz, Morriss, & McDougle, 2015). Applied behavioral analysis suggests first determining these functions ("triggers"), through tracking behavior with functional assessments (FA), before developing a management plan (Goodwin et al., 2014; Kirby, Boyd, Williams, Faldowski, & Baranek, 2016; Williams, Johnson, & Sukhodolsky, 2005). Clinicians can either complete this assessment in the clinic, or train parents/guardians and/or educators to perform FA in the home or classroom (i.e., through role playing, watching videos and modeling). In completing an assessment, detailed events before and after SIB are recorded and described, as well as any other potentially relevant information about the child or surroundings (Williams et al., 2005). Clinicians can then infer triggers from FA results, and develop a management plan to remove the trigger or redirect, replace or extinguish the SIB (Goodwin et al., 2014; Kirby et al., 2016; Williams et al., 2005).

However, there are several challenges associated with FA. SIB can persist and change functions as the individual with SIB ages, and can present differently depending on context (Kirby et al., 2016). These challenges require additional FAs in several contexts and at multiple times, as FA results can otherwise have low ecological validity (Dunlap, Newton, Fox, Benito, & Vaughn, 2001; Kirby et al., 2016). When conducted across environments, FAs can have poor consistency among observers with varying levels of experience and unstandardized protocols (Allen & Warzak, 2000; Dracobly, Dozier, Briggs, & Juanico, 2018; Marcu et al., 2013). Results may also suffer from low accuracy if there are competing demands (e.g., other children or responsibilities), or if SIB descriptions are recalled after the fact (Marcu et al., 2013; Tarbox et al., 2009; Williams et al., 2005). Further research is needed to address these limitations of FA, in particular through developing an efficient, valid, and accurate SIB monitoring method.

In addition to challenges with monitoring efficiency and data quality, FA also presents difficulties for the parents/guardians and educators ("caregivers") conducting the analyses. For example, caregivers reported sustained levels of stress when managing a high-risk behavior (i.e., SIB), and commonly disagreed with other caregivers about behavioral severity (Lecavalier, Leone, & Wiltz, 2006). FA requires high compliance and tedious efforts to track the details described above, which may deter caregivers from completing FA while concurrently trying to control the behavior (Hodgetts, Nicholas, & Zwaigenbaum, 2013). Behavioral monitoring could address these obstacles of compliance and consistency when performing FA. An automatic SIB detection method, ideally, should require limited input from caregivers, and objective metrics should be generated to quantify SIB severity across caregivers. Such new monitoring methods could address the aforementioned caregiver needs, though specific needs regarding technology for SIB should be further evaluated.

Caregiver experiences with SIB management and related technology should be considered before developing smart SIB detection. Technology development, in the design process, typically excludes users with disabilities and those who care for people with them (Krahn, Klein Walker, & Correa-De-Araujo, 2015; Liptak et al., 2008; Or & Karsh, 2009; Soares, Vannest, & Harrison, 2009; Vohra, Madhavan, Sambamoorthi, & St Peter, 2013). Such a limitation contributes to the target user (caregivers and individuals with SIB, for this research) being less likely to adopt the technology. This limitation applies to ASD communities as well. Few studies have involved caregivers of children with ASD or children with ASD when developing technology, which has been primarily limited to software (e.g., Putnam & Chong, 2008) or robotics (e.g., Conti, Di Nuovo, Buono, & Di Nuovo, 2017), with no reported studies related to sensing technology. Though prior work (Hodgetts et al., 2013) has examined family experiences with ASD and SIB, it did not evaluate technological needs. As emphasized in previous models of technology acceptance (Giesbrecht, 2013; Odom et al., 2015), however, user perspectives should be factored into the design and implementation of technology to encourage technology adoption. Caregiver perspectives should thus be included in initial stages of design, and a user-centered perspective should be embraced, by evaluating caregiver and child needs and translating them into design requirements (Karsh, Weinger, Abbott, & Wears, 2010).

Design requirements can inform technology selection to reflect caregiver needs. Behavioral monitoring technology is rapidly expanding, and includes both nonwearable and wearable options. Nonwearable sensors have been used in interventions, though mostly limited to video observation (Bellini & Akullian, 2007; Kirby et al., 2016; Schaeffer, Hamilton, & Johnson, 2016); such observation, though, invades privacy and requires tedious retrospective coding with clear potential for errors (Kirby et al., 2016). Previous work has examined the accuracy of a depth camera in monitoring stereotypical motor movements (SMM; e.g., hand-flapping or rocking) in individuals with ASD, yet classification yielded a high false positive rate (Goncalves, Rodrigues, Costa, & Soares, 2012). Wearable systems (e.g., accelerometers and inertial measurement units), however, have been used to measure stereotypy in people with ASD with up to 97% accuracy (Goodwin, Intille, Velicer, & Groden, 2008; Rad, Furlanello, & Kessler, 2016). Such systems can be integrated into healthcare (Coronato, De Pietro, & Paragliola, 2014), as shown by previous work that implemented a wrist-worn accelerometer-based activity tracking system into a hospital for use in clinical evaluations (Coronato et al., 2014). Accelerometers have also been incorporated into cellphones and watches, and these platforms also have potential to monitor self-injurious behavior. With tactile sensitivities common among people with ASD and repetitive behavioral impairments, wearable systems could pose substantial challenges for adherence if requiring skin contact. If designed with this challenge in mind, though, wearable technology could provide novel treatment methods involving feedback to the child, and could offer solutions that easily transfer into other environments. Further examination of such technology is thus required before integrating sensors into SIB monitoring.

While motion data can be captured using the sensing technology described above, a major challenge involves identifying and labeling the behaviors of interest (e.g., SIB) from these data for further analysis. Although clinicians can annotate motion data, automatic detection and prediction could make this process more efficient and accurate. Classification methods, when paired with technology, have the potential to complete analyses with minimal (or no) user input. Indeed, earlier findings support the feasibility of applying machine learning to classify movements in ASD with up to 99% accuracy (Coronato et al., 2014; Goodwin et al., 2014; Rad et al., 2015; Rad et al., 2016). However, this work was limited to detecting SMM, without a focus on SIB. SIB has rarely been included in prior classification studies. Two studies have been reported that used actors to imitate aggressive behaviors, such as SIB (Coronato et al., 2014; Plötz et al., 2012). When classifying SIB from one child with models that were trained with simulated data, classification accuracy was only on the order of 60-70% (Plötz et al., 2012). This classification performance improved using physiological sensor data when classifying episodes of aggression towards others, though the dataset did not include SIB (Ozdenizci et al., 2018). Physiological sensors may require skin contact, however, which could be a barrier to compliance for children with tactile sensitivity. Further research is thus needed to identify accurate and efficient classification methods for SIB data from accepted sensors.

Additional challenges arise when analyzing SIB data, which is both complex and variable. There is a diverse range of presentations within and between individuals who have the variable diagnosis of ASD (Amaral, Schumann, & Nordahl, 2008; Bone et al., 2015; Gowen & Hamilton, 2013). Models that are trained for each individual ("individual-level models") can provide high accuracy in detecting SIB, as demonstrated earlier (Albinali, Goodwin, & Intille, 2011; M. Goodwin et al., 2014; Goodwin, Intille, Albinali, & Velicer, 2011). However, efficiency may suffer in exchange for this high accuracy. Models encompassing a group of children and SIB types ("group-level models") could provide generalizable and efficient tracking for caregivers of multiple children with ASD and SIB. Prior work has commonly focused analyses at the individual level (Goncalves et al., 2012; Plötz et al., 2012), though, and found a loss in performance when methods were generalized to the group level, as in Ozdenizci et al., (2018). Therefore, both individual- and group-level models should be explored.

Movement characteristics ("features") could help overcome this poorer accuracy in group-level models, particularly through incorporating movement variability. Research examining SIB through a movement perspective is limited, though this approach could be fruitful given highly variable nature of SIB. Two different types of variability, linear and nonlinear, may be useful to capture aspects of SIB (Bodfish, Parker, Lewis, Sprague, & Newell, 2001; Robertson, Caldwell, Hamill, Kamen, & Whittlesey, 2013). Linear features (e.g., standard deviation) are derived from descriptive methods (Stergiou, 2004), and describe the total variability of a system (Robertson et al., 2013). Nonlinear measures of variability (e.g., entropic and fractal measures) are founded in chaos theory, and quantify the evolution of movement over a period of time (Stergiou, 2004). Previous studies of motion related to ASD have focused on simplistic or linear predictors (Min, Tewfik, Kim, & Menard, 2009); Coronato et al., 2014), yet nonlinear features of SIB could also be of use. In particular, nonlinear variability in motor control reflects the presence of pathology (Hausdorff et al., 2003; Robertson et al., 2013; Stergiou, 2004), including an ASD diagnosis (Bodfish et al., 2001; Fournier, Amano, Radonovich, Bleser, & Hass, 2014). Similarly, nonlinear measures may aid in predicting the presence of pathological behavior (SIB). Research in degenerative diseases has also found that nonlinear measures indicate the progression of a condition (Kalron, 2016; Socie, Motl, & Sosnoff, 2014). Nonlinear analyses may therefore detect the progression from non-SIB to a state of approaching or occurring SIB, supporting the long-term goal of early SIB prediction.

In summary, complex and variable SIB presents a challenging classification problem that could be addressed with emerging technological methods and advanced features. Such methods have clear potential to increase the efficiency and validity of SIB management. To explore this potential, our immediate goal was to develop and evaluate methods to classify SIB. Our long-term goal was to provide a remote monitoring system that can assist caregivers and clinicians to monitor treatment progress, as well as help children with ASD themselves, such as through a self-notification system that indicates elevated emotions. Three specific aims were completed in support of this long-term goal.

Specific Aim 1: User Perspectives**:** Assess the acceptance of potential SIB monitoring technology among children with ASD and direct caregivers (i.e., teachers, parents/guardians). Use interviews and focus groups to identify challenges inherent to traditional SIB management. Derive design considerations through qualitative content analysis.

Specific Aim 2: SIB Detection and Classification**:** Quantify the potential to identify SIB using natural movements of users with ASD, collected using wearable accelerometers (the technological approach most acceptable from Aim 1). Evaluate the performance of diverse machine learning algorithms for detecting SIB, at both the individual and group levels.

Specific Aim 3: SIB Identification with Nonlinear Predictors**:** Detect SIB using movement features and a group-level classifier. Implement a model with multiple levels to account for variability, and include nonlinear measures of motor variability as predictors of SIB. Determine if changes in variability can be used to predict SIB episodes.

Overall, this research was designed to support the future development of a smart monitoring system for SIB. Two sets of data were collected to support the above aims, including qualitative perspectives (Aim 1) and movement data containing SIB episodes for modeling approaches (Aims 2 and 3). The following dissertation reviews three studies that address the specific aims. Chapter 2 offers design considerations for SIB monitoring technology. Design guidelines were created from caregiver reports of experiences with SIB, management, and technology, and can be used to inform future technology design for individuals with ASD and caregivers. Chapter 3 examines the use of the technology accepted in Aim 1 (wearable accelerometers), and applies several machine learning algorithms to detect SIB and classify SIB types. Chapter 4 describes a group-level classifier using feature selection and reduction techniques, which also evaluated the merits of using nonlinear measures of variability as predictors of SIB. Future research is further discussed in Chapter 5, which reviews and integrates the outcomes of these three studies and provides suggestions for future work.

References

Albinali, F., Goodwin, M. S., & Intille, S. (2011). Detecting stereotypical motor movements in the classroom using accelerometry and pattern recognition algorithms. *Pervasive and Mobile Computing, 8*(1), 103. doi:10.1016/j.pmcj.2011.04.006

Allen, K. D., & Warzak, W. J. (2000). The problem of parental nonadherence in clinical behavior analysis: Effective treatment is not enough. *Journal of Applied Behavior Analysis, 33*(3), 373-391. doi:10.1901/jaba.2000.33-373

Amaral, D. G., Schumann, C. M., & Nordahl, C. W. (2008). Neuroanatomy of autism. *Trends in Neurosciences, 31*(3), 137-145. doi:https://doi.org/10.1016/j.tins.2007.12.005

APA. (2013). *Dsm 5*: American Psychiatric Association.

Baio, J., Wiggins, L., Christensen, D. L., Maenner, M. J., Daniels, J., Warren, Z., . . . White, T. (2018). Prevalence of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 sites, united states, 2014. *Mmwr Surveillance Summaries, 67*(6), 1.

Bellini, S., & Akullian, J. (2007). A meta-analysis of video modeling and video self-modeling interventions for children and adolescents with autism spectrum disorders. *Exceptional Children, 73*(3), 264-287.

Bodfish, J. W., Parker, D. E., Lewis, M. H., Sprague, R. L., & Newell, K. M. (2001). Stereotypy and motor control: Differences in the postural stability dynamics of persons with stereotyped and dyskinetic movement disorders. *American Journal on Mental Retardation, 106*(2), 123-134.

Bone, D., Goodwin, M. S., Black, M. P., Lee, C.-C., Audhkhasi, K., & Narayanan, S. (2015). Applying machine learning to facilitate autism diagnostics: Pitfalls and promises. *Journal of Autism and Developmental Disorders, 45*(5), 1121-1136. doi:10.1007/s10803-014-2268-6

Conti, D., Di Nuovo, S., Buono, S., & Di Nuovo, A. (2017). Robots in education and care of children with developmental disabilities: A study on acceptance by experienced and future professionals. *International Journal of Social Robotics, 9*(1), 51-62. doi:10.1007/s12369-016-0359-6

Coronato, A., De Pietro, G., & Paragliola, G. (2014). A situation-aware system for the detection of motion disorders of patients with autism spectrum disorders. *Expert Systems with Applications, 41*(17), 7868-7877.

Dawson, G., Rogers, S., Munson, J., Smith, M., Winter, J., Greenson, J., . . . Varley, J. (2010). Randomized, controlled trial of an intervention for toddlers with autism: The early start denver model. *Pediatrics, 125*(1), e17-e23.

Dracobly, J. D., Dozier, C. L., Briggs, A. M., & Juanico, J. F. (2018). Reliability and validity of indirect assessment outcomes: Experts versus caregivers. *Learning and Motivation, 62*, 77-90. doi:https://doi.org/10.1016/j.lmot.2017.02.007

Dunlap, G., Newton, J. S., Fox, L., Benito, N., & Vaughn, B. (2001). Family involvement in functional assessment and positive behavior support. *Focus on autism and other developmental disabilities, 16*(4), 215-221.

Fournier, K. A., Amano, S., Radonovich, K. J., Bleser, T. M., & Hass, C. J. (2014). Decreased dynamical complexity during quiet stance in children with autism spectrum disorders. *Gait & Posture, 39*(1), 420-423. doi:10.1016/j.gaitpost.2013.08.016

Giesbrecht, E. (2013). Application of the human activity assistive technology model for occupational therapy research. *Australian Occupational Therapy Journal, 60*(4), 230-240.

Goncalves, N., Rodrigues, J. L., Costa, S., & Soares, F. (2012). *Preliminary study on determining stereotypical motor movements*, United States.

Goncalves, N., Rodrigues, J. L., Costa, S., & Soares, F. (2012, 2012). *Automatic detection of stereotyped hand flapping movements: Two different approaches*.

Goodwin, M., Haghighi, M., Tang, Q., Akcakaya, M., Erdogmus, D., & Intille, S. (2014, 2014). *Moving towards a real-time system for automatically recognizing stereotypical motor movements in individuals on the autism spectrum using wireless accelerometry*.

Goodwin, M. S., Intille, S. S., Albinali, F., & Velicer, W. F. (2011). Automated detection of stereotypical motor movements. *Journal of Autism and Developmental Disorders, 41*(6), 770-782. doi:10.1007/s10803-010-1102-z

Goodwin, M. S., Intille, S. S., Velicer, W. F., & Groden, J. (2008). *Sensor-enabled detection of stereotypical motor movements in persons with autism spectrum disorder.* Paper presented at the Proceedings of the 7th international conference on Interaction design and children.

Gowen, E., & Hamilton, A. (2013). Motor abilities in autism: A review using a computational context. *Journal of Autism and Developmental Disorders, 43*(2), 323. doi:10.1007/sl0803-012-1574-0

Hausdorff, J., Schaafsma, J., Balash, Y., Bartels, A., Gurevich, T., & Giladi, N. (2003). Impaired regulation of stride variability in parkinson's disease subjects with freezing of gait. *Experimental Brain Research, 149*(2), 187-194.

Hodgetts, S., Nicholas, D., & Zwaigenbaum, L. (2013). Home sweet home? Families' experiences with aggression in children with autism spectrum disorders. *Focus on autism and other developmental disabilities, 28*(3), 166-174. doi:doi:10.1177/1088357612472932

Iwata, B. A., Pace, G. M., Dorsey, M. F., Zarcone, J. R., Vollmer, T. R., Smith, R. G., . . . Willis, K. D. (1994). The functions of self-injurious behavior: An experimental-epidemiological analysis. *27*(2), 215-240. doi:doi:10.1901/jaba.1994.27-215

Kalron, A. (2016). Gait variability across the disability spectrum in people with multiple sclerosis. *Journal of the Neurological Sciences, 361*, 1-6. doi:http://dx.doi.org/10.1016/j.jns.2015.12.012

Karsh, B.-T., Weinger, M. B., Abbott, P. A., & Wears, R. L. (2010). Health information technology: Fallacies and sober realities. *Journal of the American Medical Informatics Association : JAMIA, 17*(6), 617-623. doi:10.1136/jamia.2010.005637

Kirby, A. V., Boyd, B. A., Williams, K. L., Faldowski, R. A., & Baranek, G. T. (2016). Sensory and repetitive behaviors among children with autism spectrum disorder at home. *Autism*, 1-13.

Krahn, G. L., Klein Walker, D., & Correa-De-Araujo, R. (2015). Persons with disabilities as an unrecognized health disparity population. *American Journal of Public Health, 105*(S2), S198-S206. doi:10.2105/AJPH.2014.302182

Lam, K. S., & Aman, M. G. (2007). The repetitive behavior scale-revised: Independent validation in individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 37*(5), 855-866.

Lecavalier, L., Leone, S., & Wiltz, J. (2006). The impact of behaviour problems on caregiver stress in young people with autism spectrum disorders. *Journal of Intellectual Disability Research, 50*(3), 172-183. doi:10.1111/j.1365-2788.2005.00732.x

Liptak, G. S., Benzoni, L. B., Mruzek, D. W., Nolan, K. W., Thingvoll, M. A., Wade, C. M., & Fryer, G. E. (2008). Disparities in diagnosis and access to health services for children with autism: Data from the national survey of children's health. *J Dev Behav Pediatr, 29*(3), 152-160. doi:10.1097/DBP.0b013e318165c7a0

Lucyshyn, J. M., Albin, R. W., Horner, R. H., Mann, J. C., Mann, J. A., & Wadsworth, G. (2007). Family implementation of positive behavior support for a child with autism longitudinal, single-case, experimental, and descriptive replication and extension. *Journal of Positive Behavior Interventions, 9*(3), 131-150.

Marcu, G., Tassini, K., Carlson, Q., Goodwyn, J., Rivkin, G., Schaefer, K. J., . . . Kiesler, S. (2013). *Why do they still use paper?: Understanding data collection and use in autism education*. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Paris, France.

McTiernan, A., Leader, G., Healy, O., & Mannion, A. (2011). Analysis of risk factors and early predictors of challenging behavior for children with autism spectrum disorder. *Research in Autism Spectrum Disorders, 5*(3), 1215-1222. doi:https://doi.org/10.1016/j.rasd.2011.01.009

Min, C.-H., Tewfik, A. H., Kim, Y., & Menard, R. (2009, 2009). *Optimal sensor location for body sensor network to detect self-stimulatory behaviors of children with autism spectrum disorder*, United States.

Minshawi, N. F., Hurwitz, S., Fodstad, J. C., Biebl, S., Morriss, D. H., & McDougle, C. J. (2014). The association between self-injurious behaviors and autism spectrum disorders. *Psychology research and behavior management, 7*, 125-136. doi:10.2147/PRBM.S44635

Minshawi, N. F., Hurwitz, S., Morriss, D., & McDougle, C. J. (2015). Multidisciplinary assessment and treatment of self-injurious behavior in autism spectrum disorder and intellectual disability: Integration of psychological and biological theory and approach. *Journal of Autism and Developmental Disorders, 45*(6), 1541-1568. doi:10.1007/s10803-014-2307-3

Murphy, O., Healy, O., & Leader, G. (2009). Risk factors for challenging behaviors among 157 children with autism spectrum disorder in ireland. *Research in Autism Spectrum Disorders, 3*(2), 474-482. doi:https://doi.org/10.1016/j.rasd.2008.09.008

Odom, S. L., Thompson, J. L., Hedges, S., Boyd, B. A., Dykstra, J. R., Duda, M. A., . . . Bord, A. (2015). Technology-aided interventions and instruction for adolescents with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 45*(12), 3805-3819. doi:10.1007/s10803-014-2320-6

Or, C. K., & Karsh, B. T. (2009). A systematic review of patient acceptance of consumer health information technology. *J Am Med Inform Assoc, 16*(4), 550-560. doi:10.1197/jamia.M2888

Ozdenizci, O., Cumpanasoiu, C., Mazefsky, C., Siegel, M., Erdogmus, D., Ioannidis, S., & Goodwin, M. S. (2018). Time-series prediction of proximal aggression onset in minimally-verbal youth with autism spectrum disorder using physiological biosignals. *arXiv preprint arXiv:1809.09948*.

Pace, G. M., Iwata, B. A., Edwards, G. L., & McCosh, K. C. (1986). Stimulus fading and transfer in the treatment of self‐restraint and self‐injurious behavior. *Journal of Applied Behavior Analysis, 19*(4), 381-389.

Plötz, T., Hammerla, N. Y., Rozga, A., Reavis, A., Call, N., & Abowd, G. D. (2012). *Automatic assessment of problem behavior in individuals with developmental disabilities.* Paper presented at the Proceedings of the 2012 ACM Conference on Ubiquitous Computing.

Putnam, C., & Chong, L. (2008). *Software and technologies designed for people with autism: What do users want?* Paper presented at the Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility.

Rad, N. M., Bizzego, A., Kia, S. M., Jurman, G., Venuti, P., & Furlanello, C. (2015). Convolutional neural network for stereotypical motor movement detection in autism. *arXiv preprint arXiv:1511.01865*.

Rad, N. M., Furlanello, C., & Kessler, F. B. (2016). Applying deep learning to stereotypical motor movement detection in autism spectrum disorders.

Rice, C. (2007). Prevalence of autism spectrum disorders--autism and developmental disabilities monitoring network, six sites, united states, 2000. *MMWR Surveill Summ, 56*(1), 1-11.

Richards, C., Oliver, C., Nelson, L., & Moss, J. (2012). Self-injurious behaviour in individuals with autism spectrum disorder and intellectual disability. *56*(5), 476-489. doi:doi:10.1111/j.1365-2788.2012.01537.x

Robertson, G., Caldwell, G., Hamill, J., Kamen, G., & Whittlesey, S. (2013). *Research methods in biomechanics, 2e*: Human Kinetics.

Rojahn, J., Matson, J. L., Lott, D., Esbensen, A. J., & Smalls, Y. (2001). The behavior problems inventory: An instrument for the assessment of self-injury, stereotyped behavior, and aggression/destruction in individuals with developmental disabilities. *Journal of Autism and Developmental Disorders, 31*(6), 577-588.

Rooker, G. W., Hausman, N. L., Fisher, A. B., Gregory, M. K., Lawell, J. L., & Hagopian, L. P. (2018). Classification of injuries observed in functional classes of self-injurious behaviour. *62*(12), 1086-1096. doi:doi:10.1111/jir.12535

Schaeffer, K. M., Hamilton, K. A., & Johnson, W. L. B. (2016). Video self-modeling interventions for students with autism spectrum disorder. *Intervention in School and Clinic*, 1053451216630281.

Soares, D. A., Vannest, K. J., & Harrison, J. (2009). Computer aided self-monitoring to increase academic production and reduce self-injurious behavior in a child with autism. *Behavioral Interventions, 24*(3), 171-183. doi:10.1002/bin.283

Socie, M. J., Motl, R. W., & Sosnoff, J. J. (2014). Examination of spatiotemporal gait parameters during the 6-min walk in individuals with multiple sclerosis. *International Journal of Rehabilitation Research, 37*(4), 311-316.

Stergiou, N. (2004). *Innovative analyses of human movement*: Human Kinetics.

Tarbox, J., Wilke, A. E., Najdowski, A. C., Findel-Pyles, R. S., Balasanyan, S., Caveney, A. C., . . . Tia, B. (2009). Comparing indirect, descriptive, and experimental functional assessments of challenging behavior in children with autism. *Journal of Developmental and Physical Disabilities, 21*(6), 493-514. doi:10.1007/s10882-009-9154-8

Vohra, R., Madhavan, S., Sambamoorthi, U., & St Peter, C. (2013). Access to services, quality of care, and family impact for children with autism, other developmental disabilities, and other mental health conditions. *Autism, 18*(7), 815-826. doi:10.1177/1362361313512902

Williams, S. K., Johnson, C., & Sukhodolsky, D. G. (2005). The role of the school psychologist in the inclusive education of school-age children with autism spectrum disorders. *Journal of School Psychology, 43*(2), 117-136.

Chapter 2: Evaluating Experiences with Self-Injurious Behavior in Autism Spectrum Disorder to Facilitate Monitoring Technology Design

Abstract

Remote monitoring may assist in managing self-injurious behavior (SIB), a pervasive concern in autism spectrum disorder (ASD). However, the perspectives of affiliated stakeholders should be considered to design effective and accepted SIB tracking methods. We examined caregiver experiences to generate design suggestions for SIB monitoring technology. Twenty-three educators and 16 parents of individuals with ASD and SIB completed interviews or focus groups, in which they discussed their needs related to monitoring SIB and associated technology use. Qualitative content analysis of participant responses revealed seven main themes associated with SIB and technology: triggers, emotional responses, SIB characteristics, management, caregiver impact, child/student impact, and sensory/technology preferences. The derived themes indicated areas of emphasis for health management technology and disability monitoring. Participants expressed the need to acquire SIB data due to underlying uncertainty about the cause of the behavior, with barriers to data collection including behavior variability, child safety, and time. If monitoring system design can address these barriers with a user-centered framework, caregivers could access SIB data and obtain support for management. Design suggestions for remote monitoring for individuals with ASD and SIB include the use of transferable and durable technology with discrete placement and simple removal. The collected stakeholder perspectives provide preliminary groundwork for an effective SIB monitoring system. Future research will use iterative design, include additional stakeholders and examine caregiver and child interactions with such systems.

Keywords: *autism spectrum disorders, non-suicidal self-injury, caregivers, behavioral monitoring, qualitative analyses*

Introduction

Autism spectrum disorder and self-injurious behavior

Autism spectrum disorder (ASD) is a pervasive developmental disability that is fairly prevalent, affecting an estimated 1 in 59 youth (Baio et al., 2018). In addition to core impairments in socialization, along with restricted interests and repetitive behaviors, a majority (68%) of children with ASD struggle with severe aggressive behaviors (Kanne & Mazurek, 2011). Aggressive episodes are one of the most dangerous behaviors present in youth with ASD, and frequently manifest as self-injurious behavior (SIB; Mahatmya, Zobel, & Valdovinos, 2008). SIB can include repetitive and rhythmic behaviors such as head banging and self-hitting (Minshawi et al., 2014), and is a leading cause of hospitalization for children with ASD (Kalb et al., 2016). Caregivers are often unable to control these episodes (Hodgetts, Nicholas, & Zwaigenbaum, 2013), especially as children mature into adolescence.

Early interventions in childhood may mitigate the challenges associated with SIB. Applied behavioral analysis suggests that clinicians first complete functional assessments (FAs) to detect triggers ("antecedents" or "precipitants") before developing a treatment plan (Goodwin et al., 2014; Kirby, Boyd, Williams, Faldowski, & Baranek, 2016; Williams, Johnson, & Sukhodolsky, 2005). Clinicians can complete FAs in the clinic, or train caregivers and/or educators to perform them. In completing FAs, detailed events preceding and following SIB are recorded and described, along with other potentially relevant information such as environmental factors (Williams et al., 2005).

Such traditional FAs are often tedious for the clinician to administer and analyze. When caregivers help complete analyses in more than one context, behavioral data are often more ecologically valid than when data are limited to one setting (Dunlap, Newton, Fox, Benito, & Vaughn, 2001). However, caregivers can face challenges with other demands on attention (e.g., other children/students or household/classroom responsibilities) and adhering to analysis protocols. Caregivers may not be able to record events while trying to safely control an SIB episode (Hodgetts et al., 2013); thus, triggers may only be able to be recalled afterwards, which may adversely affect accuracy (Williams et al., 2005). Remote monitoring could mitigate such limitations and challenges, and thereby facilitate a more comprehensive and objective approach to FA. Continuing technological developments have the potential to both advance traditional management approaches for ASD and provide new options to forecast SIB onset (Goodwin, Intille, Albinali, & Velicer, 2011; Plötz et al., 2012).

Monitoring technology for ASD

Sensors that record health data (Beckham, Greene, & Meltzer-Brody, 2013; Chen, Janz, Zhu, & Brychta, 2012; Tavares & Oliveira, 2016) support evaluations and management beyond the clinic, and may advance traditional approaches such as FA for ASD. Remote monitoring technology with such sensors could provide an alternative management method for patients with limited in-person care options (Krahn, Klein Walker, & Correa-De-Araujo, 2015; Liptak et al., 2008; Vohra, Madhavan, Sambamoorthi, & St Peter, 2013). Continuously collected and accessible data could inform decisions about the time-sensitive injury risk to the patient or caregiver (Banaee, Ahmed, & Loutfi, 2013; Bates, Saria, Ohno-Machado, Shah, & Escobar, 2014; Patel, Park, Bonato, Chan, & Rodgers, 2012; Rantz et al., 2015). Technology could also enable people with ASD to self-monitor behaviors (Xin, Sheppard, & Brown, 2017), offer warnings of SIB onset with management suggestions, or help caregivers understand behavioral trends of their children (Kientz, Hayes, Westeyn, Starner, & Abowd, 2007) to support a suitable intervention. Though not specific to SIB, prior authors have concluded that technology in interventions is feasible and effective for ASD (Daniels et al., 2018; Fletcher-Watson, 2014; Odom et al., 2015), and that personal characteristics, activities performed during use, and device attributes (e.g., usability) should be considered when designing technology-based interventions (Odom et al., 2015). Several studies have also investigated different data collection methods and methods to predict ASD-related events (Goncalves, Rodrigues, Costa, & Soares, 2012; Goodwin et al., 2011; Min, Tewfik, Kim, & Menard, 2009), results of which at least the feasibility of remote monitoring with available technology. However, detected events are not specific to SIB and, to our knowledge, such remote monitoring for SIB has yet to be explored.

Care outside of the clinic may change the role of patients in health management, which suggests the importance of informatics research in home healthcare. When developing a remote monitoring system, usability and accessibility should be of high priority from ideation through deployment (Goldberg et al., 2011). However, previous technological developments (Coronato, De Pietro, & Paragliola, 2014; Goncalves et al., 2012) and evaluations of management technology often exclude perspectives of those with disabilities or those who care for people with them (Krahn et al., 2015; Liptak et al., 2008; Or & Karsh, 2009; Soares, Vannest, & Harrison, 2009; Vanderheiden & Jordan, 2012; Vohra et al., 2013); neglecting users such as caregivers during the design process may lead to inaccessible or undesired products and misused (or unused) innovations. Though prior work has explored family experiences with SIB (Hodgetts et al., 2013), it did not include technology use and was limited to familial caregivers. In contrast, a user-centered framework encourages early inclusion of stakeholders in the design process for smooth integration and adoption into everyday life (Goldberg et al., 2011; Gordon, Henderson, Holmes, Wolters, & Bennett, 2016; Or & Karsh, 2009; Panchanathan & McDaniel, 2015; Veryzer & Borja de Mozota, 2005). Before remote monitoring can effectively support its users, there is a need for informatics research to evaluate technology acceptance while considering social and cultural influences, and specific disease or disability-related barriers (Archer, Keshavjee, Demers, & Lee, 2014; Eysenbach & Jadad, 2001; Gordon et al., 2016; Hardisty et al., 2011; Karsh, Weinger, Abbott, & Wears, 2010).

Caregiver inclusion

Caregiver involvement during interventions can encourage management effectiveness and support technology use for people with cognitive disabilities (Archer et al., 2014). Caregivers, such as parents or educators, are critical contributors to the development of monitoring methods, especially given their potential interactions with health technology (Pandolfe, Wright, Slack, & Safran, 2018). High caregiver stress may affect these interactions when managing a high-risk behavior (Lecavalier, Leone, & Wiltz, 2006). SIB affects stress differently among parents and educators (Lecavalier et al., 2006), though prior work in the context of ASD or SIB has emphasized one specific caregiver (e.g., the parent; Ingersoll & Berger, 2015). As proposed in the Human Activity Assistive Technology Model (Giesbrecht, 2013; Odom et al., 2015), the definition of stakeholders in technology design should include caregivers supporting the intervention on a day-to-day basis, to capture context of use. Reflecting this model, an informatics approach should account for both patient and caregiver perspectives and activities. Designers can align new technology with current circumstances by including caregiver experiences in the design process, specifically considering the context of the deployed technology (Valdez, Holden, Novak, & Veinot, 2014). Thus, perspectives of both educators and parents, with whom the children typically spend time, should be considered in early stages of design to enhance technology acceptance and adoption (Panchanathan & McDaniel, 2015; Veryzer & Borja de Mozota, 2005). Involving users early in the design process is particularly important to assure that specific needs of individuals with disabilities are being acknowledged and addressed (Vanderheiden & Jordan, 2012).

Remote monitoring technology could help in managing SIB, yet such technology may not be adopted unless caregiver perspectives are considered during the technology design process. The present study examined parent and educator experiences with tracking SIB in ASD, and specifically evaluated needs related to monitoring technology design. We embraced user-centered design as applied to consumer health (Goldberg et al., 2011), specifically evaluating caregiver and patient needs regarding health technology for SIB and translating them into design requirements (Karsh et al., 2010). Design requirements were compiled, for future use in developing a monitoring method that automatically detects and tracks SIB.

Materials and methods

Sample and setting

Thirty-nine volunteers participated in the study, who self-identified as either parents/guardians (hereafter "parents") or educators of children/students with ASD and SIB participated in this study (Table 1). The targeted age range was 2-30 years, which was selected based on prior work (Rattaz, Michelon, & Baghdadli, 2015). Although aggression is more prevalent in children under nine years old (Kanne & Mazurek, 2011), it often persists as the child ages (Richards, Moss, Nelson, & Oliver, 2016). Parent participants discussed students whose ages ranged from 6-26 years (M = 14.1 years, SD = 6.7 years, 12 sons, 3 daughters), whereas educators discussed students whose ages were estimated as 3-22 years. Parents were asked to rate the degree to which SIB presents a challenging problem (1 = "not", 2 = "rarely", 3 = "sometimes" and "4 = "always"); these ratings ranged from 1-3, with the mean = 1.8 and mode = 2. Data collection emphasized associated caregivers (parents and educators), because many individuals with SIB were expected to be minimally verbal or “pre-verbal” based on our pilot data and prior work (Richards et al., 2016). Educator roles varied widely, and included teachers and programmatic or administrative roles in a school, often with ABA training.

Table 1. Summary information on parents and educators (M = mean, SD = standard deviation). Annual income was reported as a range (MLow = mean lower value of range, MHigh = mean higher value of range). NA =Not applicable.

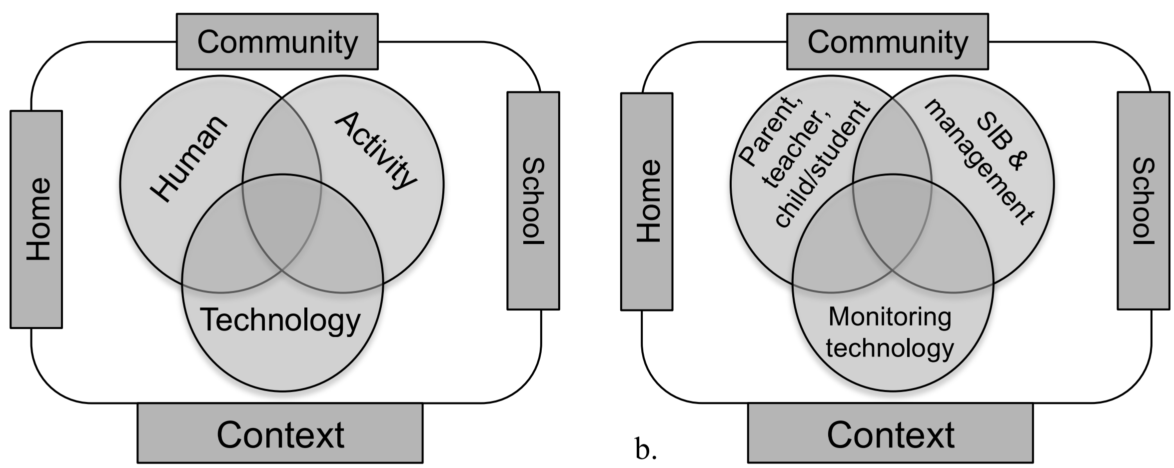
Participants	Gender	Age	Years of Experience	Focus Groups (#)
Parents	2 fathers, 14 mothers	31-62 years old (M = 45.1, SD = 8.1)	NA	2 groups (2 and 3 people)
Educators	8 male, 15 female	22-46 years old (M = 31.1, SD = 7.8)	2-24 years (M = 7.6, SD = 6.2 years)	2 groups (7 and 10 people)

Participants were recruited from the Mid-Atlantic United States, and included both rural and urban areas. Recruitment occurred at a private school offering services for ASD and through researcher social networks. All adult participants provided informed consent before any data collection, and experimental procedures were approved by the University of Virginia Institutional Review Board for Social and Behavioral Sciences and the Virginia Tech Institutional Review Board. Participants were provided with $20 gift cards and travel compensation.

Data collection

Interviews and focus groups were done concurrently to allow for triangulation methods of qualitative research. Individual interviews were conducted to allow for variability that may decrease with larger groups, and required ~1 hour. Focus groups were also included to encourage conversation among fellow caregivers (parents and educators), with a typical duration of 1.5

hours. Remote interviews occurred over the phone, whereas in-person interviews and focus groups occurred in a private conference room in a public location. Focus groups for parents were conducted separately from educator groups. An initial set of questions was developed for both the interviews and focus groups. These were open-ended and broad, to capture a wide range of responses (Hsieh & Shannon, 2005), and built on prior work related to aggression and ASD (Hodgetts et al., 2013). Questions reflected the Human Activity Assistive Technology Model (HAAT, as described in Giesbrecht, 2013; Odom et al., 2015, **Figure 1a**) applied to ASD and SIB **(Figure 1b)**, specifically addressing the following topics: child characteristics, activities of daily living and concurrent SIB, context, and monitoring technology. The HAAT model also emphasizes the importance of the context of technology (e.g., physical environment and location) and the social and cultural impacts (Giesbrecht, 2013); these aspects were considered both in the inclusion of caregivers from home and school, and in the discussion questions (e.g., Question 7, **Appendix A**). Questions spanned topics of SIB and its management, current and projected use of monitoring technology, and related benefits and challenges (e.g., Question 4c, **Appendix A).**



a. b.

Figure 1. a) Modified HAAT framework for people with ASD (adapted from Odom et al., 2015). b) HAAT framework applied to monitoring SIB to generate discussion prompts.

All but three questions remained the same for parents and educators, with adaptations made for appropriate settings (home or school) and terminology (family or students): 1) parents were further probed about the impact of SIB in their personal lives; 2) parents were asked in detail about specific preferences or aversions of the described child; and 3) educators were asked to describe the children they taught more generally, in terms of perceived diagnosis severity, to gauge their exposure to SIB. Modifications were also made when asking about a child vs. students. For example, parents were asked about the range of behaviors they experienced with their child, which differed from the questions posed to educators about the range of behaviors presented throughout their years of teaching students. Subject matter experts, including a parent, an educator, and a clinician, evaluated the initial interview questions to ensure that the questions captured relevant discussion points and comprehensively covered important characteristics of SIB and monitoring technology. Demographic questionnaires were distributed after interview or focus group completion for participant convenience, and all participants except for four educators completed the questionnaire. Feedback was also provided on the use of preferred terminology.

After posing questions during in-person interviews or focus groups, the moderator (KCG) introduced examples of monitoring technology and followed with questions to identify disparities between projected and actual needs once experiencing the technology. Included technology represented examples of available devices for remote data collection that have been explored specifically for ASD: 1) accelerometers, representing wearable devices such as smart watches (Goodwin, Intille, Velicer, & Groden, 2008), and 2) a combined depth and video camera, representing non-wearable devices (Goncalves et al., 2012; Kirby et al., 2016). Technology introduction was prefaced by noting that the involved researchers had no affiliation with, nor invested interest in, any particular technology, and emphasized that the two devices were only examples of diverse monitoring options. Note that the accelerometer was removed from its wristband to encourage discussion about other potential attachment locations. Both technology examples were distributed to focus group participants with a brief description of placement (on the body or in the environment). Questions posed about technology were repeated after technology introduction (Question 11a, **Appendix A**). All responses were audio recorded for transcription.

The initial set of questions was used in the first three parent and first three educator interviews. Analysis of results from this first stage led to a second version of interview questions to capture additional discussion topics, specifically related to technology (e.g., incorporating non-digital technology), child sensory preferences, and the wide variety of SIB definitions (see **Appendix A** for revised version). Questions were again checked by experts on the subject, and remained consistent across the remaining interviews and focus groups.

Data analysis

Results from interviews and focus groups were professionally transcribed and then reviewed by the research team prior to analysis. After the first six interviews were transcribed, four members of the team repeatedly read through the text to examine the dataset as a whole (Hsieh & Shannon, 2005). Data analysis was informed by qualitative content analysis methods (Graneheim & Lundman, 2004). Specifically, the analysis was based on Hsieh and Shannon's (2005) procedure for conventional content analysis was used to generate descriptive categories, but modified to organize results for the design community (Valdez et al., 2014). NVivo® was used to facilitate this analysis (QSR). Through discussions and consensus building, a coding manual was developed to promote consistent coding of the remaining data. The codebook included documentation of labels, definitions, and relevant examples for inclusion and exclusion in each main theme and sub-theme, or "category" (Saldaña, 2009). Themes and categories were derived directly from the data, rather than from predetermined topics, to preserve the original interview content (Hsieh & Shannon, 2005). Simultaneous coding allowed multiple meanings to be captured if several were represented within one textual sample (Saldaña, 2009). Theoretical saturation was reached after six interviews (three parent interviews and three educator interviews), though there were only minimal changes after the first three parent interviews.

Results

Main themes

Participant responses revealed seven main themes. Six themes specifically described SIB and included: (1) triggers, (2) emotional responses, (3) SIB characteristics, (4) management, (5) child/student impact, and (6) caregiver impact. Responses involving current technology motivated the theme of (7) sensory/technology preferences. **Table A1** (**Appendix A**) provides themes, subthemes, and supporting quotes for the seven main themes, which are summarized below in no particular order:

- **Triggers:** Potential antecedents (events/conditions) that may cause SIB, and can be variable within and across children/students. Parents often discussed triggers in terms of uncertainty or hope to find the trigger.
- **Emotional Response:** Reactions of the person with SIB, possibly after one of the triggers above, that lead to or occur during SIB.
- **SIB Characteristics:** Descriptions of hurtful or harmful actions the person with SIB takes towards him/herself that could or do lead to injury.
- **Management Approaches:** Methods or approaches the self or others use or have used to change one or more aspects of SIB.
- **Child/Student Impact:** Meta-level effects of SIB on people with SIB.
- **Caregiver Impact:** Meta-level effects of SIB on parents/guardians, family and educators of people with SIB.
- **Sensory/Technology Preferences:** Technology used by the child/student or by the caregiver for the child/student (e.g., SIB management material); includes external object or event that is indirectly (sensory stimuli) or directly (technology itself) related to technology that evokes a reaction from the child, causing him/her to draw near (sensory seeking) or withdraw from (sensory avoidance) that experience.

Underlying themes

Data were also cross-coded with two underlying themes, uncertainty and state of experience (hypothetical versus existing experiences). These underlying themes were evident throughout all seven main themes, and capture projected needs and concerns throughout each theme. Feelings of uncertainty often appeared to motivate participant comments about SIB definitions and origins (*"So it's hard to determine was that an SIB or are you just stimming out to the music right now?"*; *"Why would someone hurt themselves like this? What is it accomplishing for him?"*). SIB-qualifying behaviors were debated among participants in focus groups, and individual interviewees commonly questioned the interviewer if the discussed behavior was considered SIB. Further, limited communication between child/student and caregivers led to caregiver speculation about SIB triggers, child/student emotions, and long-term impact. Participants described experiences from the past and then present, as well as anticipated experiences. For example, participants included hypothetical statements about data collection systems (*"Well, you* *would love to monitor it everywhere. I mean, you would love to monitor it at home and in every class"*) and hopes for future technology to collect data on SIB (*"Along those lines, if you could data mine the original data and take not just frequency count, every gray area, you could get data on it").*

Parent/educator differences

Though parents and educators shared similar SIB experiences with the same pervasive themes in **Appendix A, Table A1**, there were several differences among the participant groups. Educators often described data collection as an integral role in their job, especially if they are part of a private school for students with ASD (*"Part of my job is to send 365 days' worth of paperwork home every quarter*"). Parents, however, noted the demands and stress levels that often prevent data collection outside of school: *"As autism moms, we don't sleep. We're on our feet like every … minute of every day. You're just so busy, there's just not time to start, stop, and write down two minutes of data …it's not practical."* Parents and educators indicated having primary knowledge of SIB exhibited either at home *or* at school, respectively, and both groups admitted this disparity was a barrier to understanding SIB. Educators also noted group observations about SIB in the school setting (*"More so the older population that I work with have engaged in the more physically abusive behaviors [than] the younger populations that I've worked with"*).

Parents and educators both described occupational impacts, but with different consequences. For example, one parent described job loss due to frequent occurrences of SIB at school, while educators discussed a change in their work environment with altered lesson plans. Financial impact was only reported in parent experiences ("*I mean, we took every penny that we had saved to get an advocate in order to get him in the placement that he needed [for SIB]*"). Relational impact was discussed among both groups of participants, yet parents tended to describe their experiences with SIB as socially isolating. Parents mentioned the difficulties encountered in public, such as when people stare or assume it is their parenting versus inquire about the disability. Community groups, such as church organizations, were mentioned as helpful support systems for parents.

Technology introduction

After being shown examples of both wearable and nonwearable technology, participants expressed reservations about child-specific use cases. For example, participants discussed possibly losing the accelerometer, and suggested adhering the accelerometer to clothing, or directly to the body to avoid problems if disrobing. Participants also indicated reservations about students noticing the depth camera in the environment, as well as its field of view. Participants indicated that a wearable device with flexible placement on the child/student, or on the caregiver (e.g., camera), could address both individual sensory concerns while remaining conspicuous.

Discussion

Parent and educator responses revealed a total of seven main themes (**Table A1, Appendix A**), with two underlying themes of uncertainty and state of experience. The first four themes encompass experiences before, during, and after SIB ("Triggers" and "Emotional Responses", "SIB Characteristics", and "Management Approaches", respectively). Our findings as related to these four themes support previous findings in several respects. Triggers identified here have also been reported in prior work, including escape from demands and cognitive rigidity (Minshawi, Hurwitz, Morriss, & McDougle, 2015). SIB types, such as head banging and self-biting, and the variable SIB duration found here, also reflect earlier findings (Minshawi et al., 2014; Minshawi et al., 2015). In contrast to prior work (Minshawi et al., 2014; Minshawi et al., 2015), though, caregivers discussed that they were uncertain as to what triggers the SIB. Caregivers also expressed uncertainty in determining whether behaviors qualify as self-injurious, which was not found in past research in the design community (Hodgetts et al., 2013). Objective health data could provide more detailed information about behavioral characteristics, which may address the uncertainty participants expressed.

Caregivers discussed management approaches after describing SIB and its potential triggers. Similar to the work of Holden et al. on chronic illness management (Holden, Valdez, Schubert, Thompson, & Hundt, 2017), our findings suggest that SIB management also depends on existing resources, caregiver engagement, and support systems. Of particular relevance for the design of monitoring systems, SIB presentation was commonly discussed as changing over time, which caregivers described as affecting management effectiveness. For example, one participant described her child self-biting at a young age, then self-scratching and head banging without self-biting at an older age. This lack of temporal consistency differs from an earlier report that SIB remained consistent, persisting decades after its onset and without significant changes in behavior type (Minshawi et al., 2014). SIB presentations were also inconsistent between individuals with ASD. This diversity in SIB may be an important potential barrier to technological interventions, which is supported by prior work (Cabibihan, Javed, Aldosari, Frazier, & Elbashir, 2017). We believe that a strength of the current study is in identifying and exploring this barrier early in the process of designing a monitoring method, specifically by obtaining perspectives from caregivers exposed to a wide range of SIB. Design considerations, summarized below, were derived from these varied SIB descriptions to encompass associated needs.

Two other themes emerged that relate to the long-term impacts of SIB on the child/student and on the caregiver. Prior work, examining SIB from national surveys, highlighted that adults with SIB suffer from social isolation (Bradley et al., 2018); our results extend this finding to youth with ASD. Also similar to earlier reports, we found that caregivers are impacted by SIB, such as through high levels of stress, relationship and financial impacts, and insufficient resources (Hodgetts et al., 2013; Lecavalier et al., 2006). Differing from prior work (Hodgetts et al., 2013; Lecavalier et al., 2006; Or & Karsh, 2009), though, we obtained parent perspectives as well as educator perspectives, and emphasized management tasks (e.g., technology use and data tracking and for remote monitoring). Our findings begin, we believe, to fill a gap in the literature regarding important factors influencing caregiver acceptance of technology. Increased caregiver acceptance implies increased caregiver support in interventions, which can promote patient adherence to remote interventions and intervention efficacy (Archer et al., 2014). Our results indicate, perhaps not surprisingly, a willingness on the part of both parents and educators to collaborate in efforts to help their child/student. They also indicated that collaboration between parents and educators supported consistent management methods, which typically resulted in more positive management outcomes. Once accepted and implemented, technology could support such collaboration by offering a unified platform for data reports and communication, which we suggest is an important area for future research.

Design considerations

The last main theme, sensory/technology preferences, was used to derive several design considerations (**Appendix A, Table A2**). Participant comments about technology were of particular interest if cross-coded with uncertainty or a hypothetical experience; these coding themes often emphasized areas of need ("*It would have to be something that is really secure in however it's being attached to the student*"; hypothetical, sensory/technology preferences). Our work, combined with earlier evidence, indicates that a wide variety of technology is used among individuals with ASD in the home and school settings (Cabibihan et al., 2017). Participants used technology identified in other work (e.g., communicative applications on smart phones or tablets; Bradley et al., 2018). However, our study contributes to a broader understanding of design options for ASD and SIB by exploring related sensory stimuli and technology use. For example, participants mentioned "soft" (non-electronic) technologies, such as "chewy bands", when discussing SIB management and sensory preferences, though these technologies were excluded in other work (Odom et al., 2015). We suggest these "soft" technologies and related sensory stimuli could be incorporated into the design of a monitoring system to encourage user acceptance, and might be used as a redirection during SIB (see **Appendix A, Table A2** for design suggestions).

Sensory preferences (e.g., pressure or vibration) could be critical aspects to include or exclude in a monitoring system, particularly if the system is designed to notify the person with SIB of his/her escalating state. Monitoring systems may be more likely adopted if integrated into commonly used technology (e.g., cellphones) and by means of a preferred attachment method (e.g., Velcro on clothing). Participants also indicated that likely attachment methods, such as a wristlet or anklet, may be tolerated after "desensitization" to common sensory aversions (Duerden et al., 2012); however, an ideal technology would not only be tolerable, but *preferred* for the child with ASD. Based on the present findings, we recommend that designers and caregivers include preferences with caution; once accustomed to a wearable or non-wearable device, participants noted that children/students could fixate on that object and even use it for sensory stimulation or self-injury (e.g., using a body-worn sensor to hit themselves, or environmental technology to aggress).

Though the technology theme relates directly to design considerations, the other six resulting themes were also interpreted in light of design considerations for SIB monitoring technology. The variety of triggers and behaviors, and the high likelihood of these changing over time, may require adaptive ("smart") monitoring systems. Our findings suggest that such systems should be customizable to accommodate SIB and ASD variability (Cabibihan et al., 2017), and to support patient and contextual variability (Valdez et al., 2014). Both parents and educators noted concerns with data collection while simultaneously managing behavior, with potential conflicts between managing child/student safety and data accuracy. Automatic detection of SIB could help alleviate responsibility from the caregiver, addressing the need to reduce workload (Holden et al., 2017; Valdez et al., 2014). Options for input, such as caregiver observations, could help caregivers account for the behavioral variability when using automatic monitoring. Our findings suggest that participants desire monitoring and management methods that translate across environments and caregivers. In an earlier report (Pandolfe et al., 2018), the authors concluded that caregiver participation in medical data systems is critical to obtaining accurate data. Thus, technology that encourages both parent and educator participation – and child/student participation if possible – could provide continuous information about SIB and across different contexts. SIB monitoring "harmony", reflecting the goal of a synchronized and patient-centered database (Pandolfe et al., 2018), requires specified caregivers and patients to have easy access to SIB data and trends. Continuous, consistent, and synchronous monitoring could provide caregivers with information about SIB across time and settings.

Strengths, limitations and future work

This study included caregivers of children with a wide age range (6-26 years old), which supports both early intervention for SIB and continued healthcare for young adults (Duerden et al., 2012; Hodgetts et al., 2013; Stahmer, Collings, & Palinkas, 2005). Though we did not target a particular gender ratio, the resulting sample discussed children who reflected the widely cited 4:1 male:female ratio children with an ASD diagnoses (Fombonne, 2003). Participants were recruited from within a relatively broad geographic region, and findings could be augmented with future work from other regions. Future research should also include a greater number of fathers. Though we collected more male perspectives than in prior work (Hodgetts et al., 2013), several of the mothers interviewed suggested their approaches to parenting often differed from the other parent involved. Since designing with stakeholders can promote long-term adoption (Gordon et al., 2016), subsequent studies should also include additional stakeholders (e.g., paid caregivers in the home and clinicians) in iterative design.

We also recommend that future work include children in design, as children may be more likely to accept technology if their perspectives are directly considered. As an extension of this study, we are conducting ongoing work using patient-centered technology and machine learning to predict SIB. Child/student perspectives will be considered in this next stage, through direct communication and through observation when the child uses technology.

Conclusions

Seven main themes emerged regarding SIB and related technology needs, which were used to offer design considerations for future monitoring technology. Caregivers were included early in the design process, which is expected to promote technology acceptance and intervention effectiveness (Panchanathan & McDaniel, 2015; Veryzer & Borja de Mozota, 2005). Once adopted, SIB monitoring could support remote care for a pervasive concern in an increasingly prevalent disability (i.e., SIB with ASD). Future work should evaluate caregiver perspectives of data sharing to develop information technology for SIB tracking.

Acknowledgements

This research was supported by a National Science Foundation Graduate Research Fellowship (to the first author) and a 4-VA Collaborative Research Grant (to RSV). However, neither agency had any involvement in data analysis, interpretation, or the decision to publish.

References

Archer, N., Keshavjee, K., Demers, C., & Lee, R. (2014). Online self-management interventions for chronically ill patients: Cognitive impairment and technology issues. *International Journal of Medical Informatics, 83*(4), 264-272. doi:https://doi.org/10.1016/j.ijmedinf.2014.01.005

Baio, J., Wiggins, L., Christensen, D. L., Maenner, M. J., Daniels, J., Warren, Z., . . . White, T. (2018). Prevalence of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 sites, united states, 2014. *Mmwr Surveillance Summaries, 67*(6), 1.

Banaee, H., Ahmed, M. U., & Loutfi, A. (2013). Data mining for wearable sensors in health monitoring systems: A review of recent trends and challenges. *Sensors (Basel), 13*(12), 17472-17500. doi:10.3390/s131217472

Bates, D. W., Saria, S., Ohno-Machado, L., Shah, A., & Escobar, G. (2014). Big data in health care: Using analytics to identify and manage high-risk and high-cost patients. *Health Affairs, 33*(7), 1123-1131.

Beckham, A. J., Greene, T. B., & Meltzer-Brody, S. (2013). A pilot study of heart rate variability biofeedback therapy in the treatment of perinatal depression on a specialized perinatal psychiatry inpatient unit. *Archives of Women's Mental Health, 16*(1), 59-65. doi:10.1007/s00737-012-0318-7

Bradley, V., Hiersteiner, D., Rotholz, D., Maloney, J., Li, H., Bonardi, A., & Bershadsky, J. (2018). Personal characteristics and outcomes of individuals with developmental disabilities who need support for self-injurious behaviour. *Journal of Intellectual Disability Research, 0*(0). doi:doi:10.1111/jir.12518

Cabibihan, J.-J., Javed, H., Aldosari, M., Frazier, T., & Elbashir, H. (2017). Sensing technologies for autism spectrum disorder screening and intervention. *Sensors, 17*(1), 46.

Chen, K. Y., Janz, K. F., Zhu, W., & Brychta, R. J. (2012). Re-defining the roles of sensors in objective physical activity monitoring. *Medicine and science in sports and exercise, 44*(1 Suppl 1), S13-S23. doi:10.1249/MSS.0b013e3182399bc8

Coronato, A., De Pietro, G., & Paragliola, G. (2014). A situation-aware system for the detection of motion disorders of patients with autism spectrum disorders. *Expert Systems with Applications, 41*(17), 7868-7877. doi:http://dx.doi.org/10.1016/j.eswa.2014.05.011

Daniels, J., Haber, N., Voss, C., Schwartz, J., Tamura, S., Fazel, A., . . . Wall, D. P. (2018). Feasibility testing of a wearable behavioral aid for social learning in children with autism. *Appl Clin Inform, 9*(1), 129-140. doi:10.1055/s-0038-1626727

Duerden, E. G., Oatley, H. K., Mak-Fan, K. M., McGrath, P. A., Taylor, M. J., Szatmari, P., & Roberts, S. W. (2012). Risk factors associated with self-injurious behaviors in children and adolescents with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 42*(11), 2460-2470. doi:10.1007/s10803-012-1497-9

Dunlap, G., Newton, J. S., Fox, L., Benito, N., & Vaughn, B. (2001). Family involvement in functional assessment and positive behavior support. *Focus on autism and other developmental disabilities, 16*(4), 215-221.

Eysenbach, G., & Jadad, A. R. (2001). Evidence-based patient choice and consumer health informatics in the internet age. *Journal of Medical Internet Research, 3*(2), e19. doi:10.2196/jmir.3.2.e19

Fletcher-Watson, S. (2014). A targeted review of computer-assisted learning for people with autism spectrum disorder: Towards a consistent methodology. *Review Journal of Autism and Developmental Disorders, 1*(2), 87-100. doi:10.1007/s40489-013-0003-4

Fombonne, E. (2003). Epidemiological surveys of autism and other pervasive developmental disorders: An update. *J Autism Dev Disord, 33*(4), 365-382.

Giesbrecht, E. (2013). Application of the human activity assistive technology model for occupational therapy research. *Australian Occupational Therapy Journal, 60*(4), 230-240.

Goldberg, L., Lide, B., Lowry, S., Massett, H. A., O'Connell, T., Preece, J., . . . Shneiderman, B. (2011). Usability and accessibility in consumer health informatics: Current trends and future challenges. *American Journal of Preventive Medicine, 40*(5, Supplement 2), S187-S197. doi:https://doi.org/10.1016/j.amepre.2011.01.009

Goncalves, N., Rodrigues, J. L., Costa, S., & Soares, F. (2012, 2012). *Automatic detection of stereotyped hand flapping movements: Two different approaches*.

Goodwin, M., Haghighi, M., Tang, Q., Akcakaya, M., Erdogmus, D., & Intille, S. (2014, 2014). *Moving towards a real-time system for automatically recognizing stereotypical motor movements in individuals on the autism spectrum using wireless accelerometry*.

Goodwin, M. S., Intille, S. S., Albinali, F., & Velicer, W. F. (2011). Automated detection of stereotypical motor movements. *Journal of Autism and Developmental Disorders, 41*(6), 770-782. doi:10.1007/s10803-010-1102-z

Goodwin, M. S., Intille, S. S., Velicer, W. F., & Groden, J. (2008). *Sensor-enabled detection of stereotypical motor movements in persons with autism spectrum disorder.* Paper presented at the Proceedings of the 7th international conference on Interaction design and children.

Gordon, M., Henderson, R., Holmes, J. H., Wolters, M. K., & Bennett, I. M. (2016). Participatory design of ehealth solutions for women from vulnerable populations with perinatal depression. *J Am Med Inform Assoc, 23*(1), 105-109. doi:10.1093/jamia/ocv109

Graneheim, U. H., & Lundman, B. (2004). Qualitative content analysis in nursing research: Concepts, procedures and measures to achieve trustworthiness. *Nurse education today, 24*(2), 105-112.

Hardisty, A. R., Peirce, S. C., Preece, A., Bolton, C. E., Conley, E. C., Gray, W. A., . . . Elwyn, G. (2011). Bridging two translation gaps: A new informatics research agenda for telemonitoring of chronic disease. *International Journal of Medical Informatics, 80*(10), 734-744. doi:https://doi.org/10.1016/j.ijmedinf.2011.07.002

Hodgetts, S., Nicholas, D., & Zwaigenbaum, L. (2013). Home sweet home? Families' experiences with aggression in children with autism spectrum disorders. *Focus on autism and other developmental disabilities, 28*(3), 166-174. doi:doi:10.1177/1088357612472932

Holden, R. J., Valdez, R. S., Schubert, C. C., Thompson, M. J., & Hundt, A. S. (2017). Macroergonomic factors in the patient work system: Examining the context of patients with chronic illness. *Ergonomics, 60*(1), 26-43.

Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative health research, 15*(9), 1277-1288.

Ingersoll, B., & Berger, N. I. (2015). Parent engagement with a telehealth-based parent-mediated intervention program for children with autism spectrum disorders: Predictors of program

use and parent outcomes. *Journal of Medical Internet Research, 17*(10), e227. doi:10.2196/jmir.4913
Kalb, L. G., Vasa, R. A., Ballard, E. D., Woods, S., Goldstein, M., & Wilcox, H. C. (2016). Epidemiology of injury-related emergency department visits in the us among youth with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 46*(8), 2756-2763. doi:10.1007/s10803-016-2820-7
Kanne, S. M., & Mazurek, M. O. (2011). Aggression in children and adolescents with asd: Prevalence and risk factors. *Journal of Autism and Developmental Disorders, 41*(7), 926-937. doi:10.1007/s10803-010-1118-4
Karsh, B.-T., Weinger, M. B., Abbott, P. A., & Wears, R. L. (2010). Health information technology: Fallacies and sober realities. *Journal of the American Medical Informatics Association : JAMIA, 17*(6), 617-623. doi:10.1136/jamia.2010.005637
Kientz, J. A., Hayes, G. R., Westeyn, T. L., Starner, T., & Abowd, G. D. (2007). Pervasive computing and autism: Assisting caregivers of children with special needs. *IEEE Pervasive Computing, 6*(1).
Kirby, A. V., Boyd, B. A., Williams, K. L., Faldowski, R. A., & Baranek, G. T. (2016). Sensory and repetitive behaviors among children with autism spectrum disorder at home. *Autism*, 1-13.
Krahn, G. L., Klein Walker, D., & Correa-De-Araujo, R. (2015). Persons with disabilities as an unrecognized health disparity population. *American Journal of Public Health, 105*(S2), S198-S206. doi:10.2105/AJPH.2014.302182
Lecavalier, L., Leone, S., & Wiltz, J. (2006). The impact of behaviour problems on caregiver stress in young people with autism spectrum disorders. *Journal of Intellectual Disability Research, 50*(3), 172-183. doi:10.1111/j.1365-2788.2005.00732.x
Liptak, G. S., Benzoni, L. B., Mruzek, D. W., Nolan, K. W., Thingvoll, M. A., Wade, C. M., & Fryer, G. E. (2008). Disparities in diagnosis and access to health services for children with autism: Data from the national survey of children's health. *J Dev Behav Pediatr, 29*(3), 152-160. doi:10.1097/DBP.0b013e318165c7a0
Mahatmya, D., Zobel, A., & Valdovinos, M. G. (2008). Treatment approaches for self-injurious behavior in individuals with autism: Behavioral and pharmacological methods. *Journal of Early and Intensive Behavior Intervention, 5*(1), 106-118. doi:10.1037/h0100413
Min, C.-H., Tewfik, A. H., Kim, Y., & Menard, R. (2009, 2009). *Optimal sensor location for body sensor network to detect self-stimulatory behaviors of children with autism spectrum disorder*, United States.
Minshawi, N. F., Hurwitz, S., Fodstad, J. C., Biebl, S., Morriss, D. H., & McDougle, C. J. (2014). The association between self-injurious behaviors and autism spectrum disorders. *Psychology research and behavior management, 7*, 125-136. doi:10.2147/PRBM.S44635
Minshawi, N. F., Hurwitz, S., Morriss, D., & McDougle, C. J. (2015). Multidisciplinary assessment and treatment of self-injurious behavior in autism spectrum disorder and intellectual disability: Integration of psychological and biological theory and approach. *Journal of Autism and Developmental Disorders, 45*(6), 1541-1568. doi:10.1007/s10803-014-2307-3
Odom, S. L., Thompson, J. L., Hedges, S., Boyd, B. A., Dykstra, J. R., Duda, M. A., . . . Bord, A. (2015). Technology-aided interventions and instruction for adolescents with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 45*(12), 3805-3819. doi:10.1007/s10803-014-2320-6

Or, C. K., & Karsh, B. T. (2009). A systematic review of patient acceptance of consumer health information technology. *J Am Med Inform Assoc, 16*(4), 550-560. doi:10.1197/jamia.M2888

Panchanathan, S., & McDaniel, T. (2015). Person-centered accessible technologies and computing solutions through interdisciplinary and integrated perspectives from disability research. *Universal Access in the Information Society, 14*(3), 415-426. doi:10.1007/s10209-014-0369-9

Pandolfe, F., Wright, A., Slack, W. V., & Safran, C. (2018). Rethinking the outpatient medication list: Increasing patient activation and education while architecting for centralization and improved medication reconciliation. *Journal of the American Medical Informatics Association, 25*(8), 1047-1053. doi:10.1093/jamia/ocy047

Patel, S., Park, H., Bonato, P., Chan, L., & Rodgers, M. (2012). A review of wearable sensors and systems with application in rehabilitation. *J Neuroeng Rehabil, 9*, 21-21. doi:10.1186/1743-0003-9-21

Plötz, T., Hammerla, N. Y., Rozga, A., Reavis, A., Call, N., & Abowd, G. D. (2012). *Automatic assessment of problem behavior in individuals with developmental disabilities.* Paper presented at the Proceedings of the 2012 ACM Conference on Ubiquitous Computing.

QSR. *QSR I. NVivo. 9 ed.*

Rantz, M. J., Skubic, M., Popescu, M., Galambos, C., Koopman, R. J., Alexander, G. L., . . . Miller, S. J. (2015). A new paradigm of technology-enabled 'vital signs' for early detection of health change for older adults. *Gerontology, 61*(3), 281-290. doi:10.1159/000366518

Rattaz, C., Michelon, C., & Baghdadli, A. (2015). Symptom severity as a risk factor for self‐injurious behaviours in adolescents with autism spectrum disorders. *Journal of Intellectual Disability Research, 59*(8), 730-740. doi:10.1111/jir.12177

Richards, C., Moss, J., Nelson, L., & Oliver, C. (2016). Persistence of self-injurious behaviour in autism spectrum disorder over 3 years: A prospective cohort study of risk markers. *J Neurodev Disord, 8*, 21. doi:10.1186/s11689-016-9153-x

Saldaña, J. (2009). *The coding manual for qualitative researchers*. Thousand Oaks, Calif;London;: Sage.

Soares, D. A., Vannest, K. J., & Harrison, J. (2009). Computer aided self-monitoring to increase academic production and reduce self-injurious behavior in a child with autism. *Behavioral Interventions, 24*(3), 171-183. doi:10.1002/bin.283

Stahmer, A. C., Collings, N. M., & Palinkas, L. A. (2005). Early intervention practices for children with autism: Descriptions from community providers. *Focus on autism and other developmental disabilities, 20*(2), 66-79. doi:doi:10.1177/10883576050200020301

Tavares, J., & Oliveira, T. (2016). Electronic health record patient portal adoption by health care consumers: An acceptance model and survey. *Journal of Medical Internet Research, 18*(3).

Valdez, R. S., Holden, R. J., Novak, L. L., & Veinot, T. C. (2014). Transforming consumer health informatics through a patient work framework: Connecting patients to context. *Journal of the American Medical Informatics Association, 22*(1), 2-10.

Vanderheiden, G. C., & Jordan, J. B. (2012). Design for people with functional limitations. *Handbook of Human Factors and Ergonomics, Fourth Edition*, 1407-1441.

Veryzer, R. W., & Borja de Mozota, B. (2005). The impact of user ‐ oriented design on new product development: An examination of fundamental relationships. *Journal of Product innovation management, 22*(2), 128-143.

Vohra, R., Madhavan, S., Sambamoorthi, U., & St Peter, C. (2013). Access to services, quality of care, and family impact for children with autism, other developmental disabilities, and other mental health conditions. *Autism, 18*(7), 815-826. doi:10.1177/1362361313512902

Williams, S. K., Johnson, C., & Sukhodolsky, D. G. (2005). The role of the school psychologist in the inclusive education of school-age children with autism spectrum disorders. *Journal of School Psychology, 43*(2), 117-136.

Xin, J. F., Sheppard, M. E., & Brown, M. (2017). Brief report: Using ipads for self-monitoring of students with autism. *Journal of Autism and Developmental Disorders, 47*(5), 1559-1567. doi:10.1007/s10803-017-3055-y

Chapter 3: Detecting and Classifying Self-Injurious Behavior in Autism Spectrum Disorder using Machine Learning Techniques

Abstract

Self-injurious behavior (SIB) is a critical problem among people with autism spectrum disorder (ASD). Traditional SIB management methods often have low ecological validity and accuracy, and place compliance demands on the caregiver. Sensing technology could address these concerns, particularly through monitoring movements, though effective classification of these movements will be required. To support an SIB monitoring system, this study evaluated several machine learning algorithms for detecting SIB among individuals and groups of individuals, and determined the performance of such methods in distinguishing between several types of SIB. Ten children with ASD and SIB and their caregivers participated in free play with interspersed tasks at a child study center, during which >200 episodes of 18 different SIB types were captured with body-worn accelerometers. In detecting SIB versus non-SIB, as well as in discriminating between multiple types of behaviors, classifiers using k-nearest neighbors and support vector machines provided the highest accuracy both for individuals (up to 99.1%) and for grouped participants (up to 94.6%). In all cases, classification efficiency was quite high, with an offline rate of >8,600 Hz (~0.1 msec/classification). These performance levels were achieved among individuals spanning a wide age range and who exhibited a diverse set of behaviors, supporting the robustness of these classification methods. Our results support the feasibility of continuous and objective SIB monitoring, which could in turn facilitate the future care of a pervasive concern in ASD.

Keywords: *autism spectrum disorder, self-injurious behaviors, activity recognition, machine learning, wearable sensors*

Introduction

Self-injurious behavior (SIB) is a leading cause of hospitalization for children with autism spectrum disorder (ASD; Kalb et al., 2016). SIB may be repetitive and rhythmic, and can include behaviors such as head banging and self-hitting (Minshawi et al., 2014). Physical damage may result (Rooker et al., 2018), including abrasions, lacerations, and contusions, especially as SIB commonly continues beyond the initial age of onset (Minshawi et al., 2014; Richards, Moss, Nelson, & Oliver, 2016; Taylor, Oliver, & Murphy, 2011). However, early interventions can help prevent severe consequences and alleviate the long-term persistence of SIB (Kurtz et al., 2003).

As suggested by applied behavioral analysis, functional assessments (FA) should be completed before an intervention to determine potential triggers of SIB (Iwata et al., 1994; Pelios, Morren, Tesch, & Axelrod, 1999; Williams, Johnson, & Sukhodolsky, 2005). In typical applications, FA requires clinicians or trained caregivers to track behavior times and the events surrounding the behavior (Williams et al., 2005). Clinicians can then interpret FA results to infer potential triggers and plan interventions. SIB, though, can present differently depending on context (Kirby, Boyd, Williams, Faldowski, & Baranek, 2016), and data from FAs may have low ecologically validity if limited to a single environment (Dunlap, Newton, Fox, Benito, & Vaughn, 2001). When FAs are performed across environments, caregivers face challenges in consistently monitoring behaviors, due to demands on their attention (e.g., other children/students or household/classroom responsibilities), and FA results can vary between clinicians and caregivers (Allen & Warzak, 2000; Dracobly, Dozier, Briggs, & Juanico, 2018; Marcu et al., 2013). Manually recorded descriptions in FAs can be unstandardized, and traditional paper and pencil approaches require tedious annotating (Dracobly et al., 2018; Marcu et al., 2013). Results may also suffer from low accuracy if the observer is inexperienced, or if SIB descriptions are recalled after the fact (Marcu et al., 2013; Tarbox et al., 2009). An ongoing need thus exists to identify efficient SIB monitoring methods, specifically to support acquiring accurate and valid behavioral data in FAs.

Monitoring technology

Recent developments in monitoring technologies could overcome the challenges of FA, particularly through providing comprehensive and objective behavioral data (Cabibihan, Javed, Aldosari, Frazier, & Elbashir, 2017; Goodwin, Intille, Albinali, & Velicer, 2011; Plötz et al., 2012).

This technology is rapidly expanding, and includes both nonwearable and wearable systems (Cabibihan et al., 2017; Zheng et al., 2014). Nonwearable technology has been used in ASD interventions, but is mainly limited to video observation (Bellini & Akullian, 2007; Kirby et al., 2016; Schaeffer, Hamilton, & Johnson, 2016; Soares, Vannest, & Harrison, 2009). For example, retrospective coding of home videos provided information on triggers for sensory stimulation and repetitive behavior, specifically stereotypical motor movements (SMM; Kirby et al., 2016); however, use of video involves tedious human coding with clear potential for errors. Past studies have also examined the accuracy of a depth camera in monitoring SMM (e.g., hand-flapping) in ASD, yet false positives were common from noise, limited field of view, and signals from interceding caregivers (Goncalves, Rodrigues, Costa, & Soares, 2012). Nonwearable systems may thus fail to provide the accuracy, efficiency, or transferability needed in an SIB monitoring system.

Wearable systems, such as accelerometers and inertial measurement units, could address the limitations of nonwearable devices, and provide the flexibility and accuracy needed to monitor SIB across multiple environments and with several caregivers (Bulling, Blanke, & Schiele, 2014; Goncalves et al., 2012). For example, accelerometers were used to record SMM, including hand-flapping and rocking in children with ASD, with decision trees yielding 82-97% accuracy using sensors on the wrist and back (Goodwin, Intille, Velicer, & Groden, 2008). Other work indicated that only one accelerometer was necessary to track both flapping and rocking, and that a sensor placed on the upper back could classify both types of SMM with accuracies of 80.5-95.5% (Min, Tewfik, Kim, & Menard, 2009). Such results suggest that repetitive behaviors can be detected without a need for invasive sensors. Coronato et al. (2014) implemented an activity tracking system using a wrist-worn accelerometer in a hospital for use in clinical evaluations of ASD, highlighting the feasibility of integration into healthcare. Recent work implies that wearables could provide sufficient data to classify even finer motor movements that could be injurious, such as scratching (Moreau et al., 2018). Wearables therefore offer a promising method to continuously collect movement data for SIB monitoring.

In addition to providing technical advantages, wearable devices could be integrated into patient care if accepted by both the individual with ASD and the caregiver. Wearables could serve as an avenue for novel management methods, such as by offering feedback to the child, and could yield solutions that transfer into other environments. Such technology will not be worn, however, if not first accepted by the wearer and other users. In our previous study on caregiver perspectives of monitoring technology for SIB, we found that people with ASD and their caregivers would likely accept wearables (Cantin-Garside et al., 2018), despite prevalent sensory sensitivity (Chen, Rodgers, & McConachie, 2009). Movement sensors, in contrast to commonly used physiological sensors, do not require direct contact with the skin and can be creatively placed to avoid the aforementioned sensitivities. Caregivers also expressed the need for a transferable solution, which can be provided by wearables (Cantin-Garside et al., 2018). Wearable sensors can stream continuous and comprehensive data to monitor SIB across environments, and were therefore used in the current study. These devices support smart monitoring through data collection, feeding the collected data as input into classifiers for SIB detection.

Activity classification

SIB monitoring, however, requires advanced classification methods to detect SIB from within complex streams of movement data, and which should yield both high efficiency (i.e., low training and classification times) and high accuracy (i.e., correctly classified SIB and non-SIB events). Commonly employed algorithms for use in human activity detection include support vector machines (SVM), discriminant analyses (DA), decision trees (DT), Naïve Bayes (nB), k-nearest neighbor (kNN), and neural networks (NN) (Kim & Nussbaum, 2014; McLeod et al., 2016; Miller, Beazer, & Hahn, 2013; Mittek, Carlson, Mora-Becerra, Psota, & Perez, 2015; Moreau et al., 2018). More detailed reviews are available elsewhere (Banaee, Ahmed, & Loutfi, 2013; Preece et al., 2009; Wu et al., 2008). Machine learning algorithms have been applied to classify movement among people with ASD. For example, NN and SVM were up to 99% accurate when detecting repetitive motor movements (Coronato et al., 2014; Goodwin et al., 2014). Previously explored classification methods may extend to SIB, particularly because previous work found a relationship between SIB occurrence and stereotypical motor movements (Minshawi et al., 2014). There is thus the potential of applying machine learning to monitor movement with ASD, although existing work has been limited to detecting stereotypy without naturally occurring SIB. Novel approaches to SIB classification should also be explored. Sparse representation classification (SRC) is a relatively new approach for activity detection, which operates under the assumption that data from the same activity will be highly correlated (Bastani, Rao, & Kong, 2016). In a study on repetitive manual handling tasks, this approach outperformed traditional methods (SVM, DA and NN) in terms of accuracy, while maintaining acceptable efficiency (Bastani, Kim, Kong, Nussbaum, & Huang, 2016). SRC classified stereotypical motor movements with high accuracy as well (Min et al., 2009); however, this method has not yet been applied to SIB, and SRC could provide an accurate and efficient online SIB monitoring.

Previous research on SIB classification has been limited in application and scope, focusing on imitated behaviors or relatively few SIB observations. For example, prior work created a classifier from trained actors to imitate aggressive behaviors, including SIB (Coronato et al., 2014; Plötz et al., 2012). One study found accuracies of 63.8-69.6%, 57.8-63.5%, and 69.4-71.6% for nB, DTs, and SVMs, respectively, when extending methods from simulated data, using actors, to a child with SIB wearing upper-body accelerometers (Plötz et al., 2012). These results, however, may not translate to natural data. Improved classification performance was achieved when integrating physiological sensors to detect the onset of aggression towards others, though SIB was not included as an activity of interest (Ozdenizci et al., 2018). Min et al. (2017) used a Hidden Markov model to classify self-stimulatory behaviors among four participants; one of the four participants also showed behavior that could be considered SIB (punching), though SIB was not specifically analyzed. This previous research supports the feasibility of monitoring ASD behaviors through wearable sensing, yet it remains to demonstrate whether classification can be effective for a range of natural SIB data.

SIB presents a complex classification problem, however, with a diverse range of presentations both within and between individuals who have the variable diagnosis of ASD (Amaral, Schumann, & Nordahl, 2008; Bone et al., 2015; Gowen & Hamilton, 2013). High levels of noise are expected when equipping children with sensors, and robust methods are required to handle sensor configurations that could differ according to acceptance (e.g., one child may prefer only upper-body sensors, as in Plötz et al., 2012). Further, the feasibility of group-level classification should also be examined, given the enhanced efficiency and generalizability, though previous work either focused analyses at the individual level (Goncalves et al., 2012; Plötz et al., 2012) or found a loss in performance when methods were generalized to the group level (Ozdenizci et al., 2018).

Innovations in monitoring technologies and machine learning algorithms have the potential to alleviate the burdens of traditional SIB monitoring and provide a platform to inform future interventions. While past efforts support the potential of classifying SIB behavior in ASD, direct evidence is currently lacking. As such, the present work assumed an exploratory approach, and included a range of classification methods. The study reported here aimed to: (1) evaluate the performance of different machine learning algorithms for SIB detection at both the individual and group levels, and (2) determine the performance of such algorithms in distinguishing between several types of SIB. Addressing these aims was considered an initial step toward creating a smart SIB monitoring system.

Materials and methods

Participants

Participants were recruited through the Virginia Tech Child Study Center and through the authors' networks. Children aged 5-14 years with a diagnosis of ASD were of specific interest, due to heightened aggression during childhood (Hill et al., 2014; Kanne & Mazurek, 2011). To ensure we could capture multiple episodes in a 1-3 hour clinical session, included children had to exhibit SIB >3/hr. Definitions of SIB vary (Lam & Aman, 2007; Rojahn, Matson, Lott, Esbensen, & Smalls, 2001); here, SIB was defined as an action that could cause harm or pain to the self and that may serve different functions. SIB excluded accidental incidents (e.g., tripping and falling) and behaviors intended for cultural or aesthetic reasons (e.g., ear piercings). SIB itself may or may not be performed with the intention to harm. It can be repetitive, and can include classic behaviors (e.g., head banging, self-hitting or biting), fine behaviors (e.g., picking or pinching), and potentially injurious body contact with objects or the environment (e.g., repeatedly hitting into walls). This definition reflects definitions and SIB types in prior work (Iwata et al., 1994; Pace, Iwata, Edwards, & McCosh, 1986; Rojahn et al., 2001). For this study, an SIB "episode" was defined as a continuous time period of a specific behavior type, of SIB or of co-occurring SIB, preceded and followed by other types of SIB or other non-SIB actions.

Caregivers were pre-screened to confirm our inclusion criteria: fluency in English and home within driving distance of the Virginia Tech Child Study Center; and child age, SIB, and diagnostic information described above. Caregivers were also asked, during pre-screening, to describe the child's SIB, typical SIB triggers, and preferred materials for sensors. We requested that caregivers bring toys for free play, toys or props affiliated with starting and stopping SIB, and preferred clothes or accessories. We also asked that caregivers dress their children in pants and/or sweatshirts with pockets for sensor placement. All adult participants provided informed consent, and qualifying children provided assent (if >7 years of age and of developmental level), before any data collection. The Virginia Tech Institutional Review Board approved all experimental procedures.

Study overview

Participants came to the noted Study Center, where the lead author, or the caregiver guided by the author, secured sensors on the child after consent (see **Table 2**). A clinical psychology doctoral candidate then confirmed ASD diagnosis using standard tools: ADOS and WASI, or Leite-R if the child was nonverbal (Corporation, 1999; Lord et al., 2012; Pugliese et al., 2015). The caregiver completed a demographic questionnaire with a list of potential triggers for SIB. Each child was then monitored during free-play, using movement sensors and video recording (see below), and by 2-3 observing researchers. Participants remained at the Study Center until either: a) >3 episodes of SIB were captured, or b) the session was terminated to avoid escalating behavior.

If sufficient SIB episodes did not occur naturally, caregivers were able to continue the session with the Standardized Observation Analogue Procedure (SOAP), which allowed them to prompt for the injurious behavior in a controlled fashion (Johnson et al., 2009). Specifically, caregivers selected common everyday demands from a list of triggers (e.g., "Hand me the toy"), or personalized demands that could promote resistance and relate to a specific SIB trigger (e.g., removing shoes and leaving socks). The caregiver issued a new demand every minute until either the target behavior occurred or the caregiver and participant elected to stop (see Johnson et al., 2009 for a more detailed review). Caregivers were instructed to manage SIB as they would at home. The psychology doctoral candidate, overseen by a clinical psychologist (SWW), supervised the process to ensure prompt intervention when needed. At the end of the session, participants were compensated for their travel and time, and children selected a toy.

Instrumentation

Tri-axial accelerometers (ActiGraph GT9X Link, www.actigraphcorp.com) were used to collect movement data (60 Hz) for the duration of each child's time at the Study Center. Based on our earlier results from caregiver focus groups, accelerometers were selected as discrete and transferable devices with flexible placement (Cantin-Garside et al., 2018). Accelerometers were applied at the wrists, waist, and/or ankles, or within pant pockets, as accepted by each child (see **Table 2**). These placement options were based, in part, on prior work showing that movement sensors on either the wrist or torso provide high accuracy in classifying activities (Min et al., 2009; Trost, Zheng, & Wong, 2014) and that added ankle sensors may be necessary to capture lower-body injurious behaviors (Plötz et al., 2012). A maximum of six accelerometers were placed on each child, and preferred materials (cotton, fleece, or sequins) were added to the sensors as he/she desired (see **Discussion** for a summary of accepted materials). Video recordings were also obtained for each child, using three Go-Pro™ cameras and an overhead camera. These recordings provided "ground truth" for labeling the accelerometer data.

Data preprocessing

Data from each channel of each accelerometer were low-pass filtered (fourth-order, Butterworth, bi-directional) with a cutoff frequency of 20 Hz. The frequency of SIB is unknown due to the lack of prior research; thus, this cutoff value was conservatively selected to preserve higher frequency components of SIB movement data. Filtered data from each channel served as input to classification methods described subsequently.

Before labeling the accelerometer data, behavioral definitions were discussed among members of our research team. The team also discussed the SIB described during prescreening by the caregiver, and the behaviors seen during and after each session. Multiple researchers annotated and discussed the video data, and two labeling schemes were created from these annotated files. The first scheme was used for the purpose of SIB detection (Aim 1): 0 = non-SIB, or "other" behaviors; and 1 = SIB. The second scheme used specific labels (0, …, 23) for each different SIB presentation that was within the definition provided above, thereby representing different SIB types (Aim 2). Video data were viewed and labeled every second. Subsequently, we grouped participant data according to SIB type (see **Table 2**). As described in the next section, both individual levels and group-level classifiers were trained, validated, and tested for both the (0,1) and the (0-23) labeling schemes (see **Figure 2** for a summary of the datasets used).

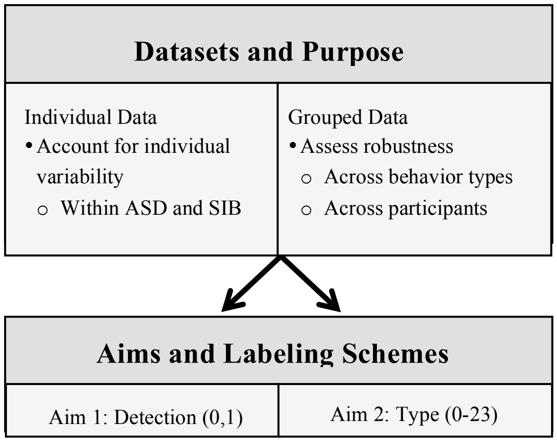


Figure 2. Datasets used to address each aim.

Classification methods and evaluation

As in other work on detecting rare events (Albinali, Goodwin, & Intille, 2009; Goodwin et al., 2014; Großekathöfer et al., 2017), SIB is considerably underrepresented when compared to other behaviors in a naturalistic dataset. Here, limited samples of SIB, together with a large class of non-SIB activity ("other"), results in a highly skewed distribution that can compromise classification performance. SIB durations lasted a minimum of about two seconds (though more subtle movements, like picking, lasted longer, >~10 seconds).

To balance these data, *N* data points (observations collected at 60 Hz) from all worn sensors were randomly selected from each label (Albinali et al., 2009; Bastani, Kim, et al., 2016; Bastani, Rao, et al., 2016; Großekathöfer et al., 2017; Rad et al., 2016). In preliminary work, we examined the accuracy and efficiency of validation results from training with 10 different *N* values, ranging from 100 to 1000 per label in equal increments (Bastani, Kim, et al., 2016; Großekathöfer et al., 2017). From these initial analyses, the number of training data points was selected as $N = 500$ for the first label scheme (0,1) and 400 for the second label scheme (0-23). These training sizes are comparable to previous work (Bastani, Kim, et al., 2016; Coronato et al., 2014), and preliminary analyses indicated that additional observations improved classification performance only marginally with larger *N* values. As discussed in Bastani, Kim, et al. (2016), training points were assumed to reflect the entire dataset, since they were randomly selected from across the duration of a given session. The selected numbers of data points for each labeling scheme were considered as representative, yet efficient, training sizes.

We developed several classification models using a range of machine learning algorithms, given the lack of relevant literature related to SIB. Specific algorithms included were SVM, kNN, SRC, DT, DA, nB, and NN. Note that training, validation, and test data were randomly selected and independent from each other. Validation data were used for model selection, and test data were used for final model evaluation (Poliker, 2006). Validation data were randomly selected and held out from the balanced set noted above, such that the ratio of training to validation data was 3:1, with training size set as the determined *N* above. Absent guidance from prior work specifically targeting SIB, we implemented a featureless approach here. A method independent of features classifies movement data in its true form and avoids selecting misrepresentative predictors of SIB (Bastani, Kim, et al., 2016; Plötz, Hammerla, & Olivier, 2011). Featureless methods also offer efficiency, with minimal preprocessing required (Bastani, Kim, et al., 2016; Bastani, Rao, et al., 2016).

Additional preliminary work was completed in which classification models were trained and validated using multiple parameter levels. More specifically, we explored: several kernels for SVM (linear, cubic, quadratic and multiple Gaussian kernels); k = 1 – 10 for kNN, as well as weighted, cubic and cosine kNN; fine, medium and course DT; linear and quadratic DA; and NN using 1-30 hidden layers. In all cases, we used MatLab R2018a (Classification Learner and Deep Learning Toolbox, on a CPU of Intel dual-core 2.6 GHz) to train, validate, and test each classification model. Methods described in Gaonkar and Davatzikos (2013) and Wright et al. (2009) were used for SRC.

Outcome measures assessing model performance included the following, derived from the validation datasets: accuracy, specificity, F-score, and efficiency (total time to classify a single data point). Models with parameters that resulted in the highest accuracy and efficiency were selected for final testing (see **Appendix B**, **Table B1** for results from this initial stage of validation). For the first labeling scheme (0, 1), assessing SIB detection, the following classifiers were employed: 1) kNN (with k = 2 or 3 at the individual/group levels); 2) SVM with Gaussian Kernel; 3) NN with 20 hidden layers, similar to prior work (Kim & Nussbaum, 2014, Bastani, Kim, et al., 2016); 4) DT with 100 splits; and 5) quadratic DA. For the second labeling scheme, assessing the ability to distinguish SIB (0-23 labels), we used the following classifiers: 1) kNN, (k = 4 or 2 at the individual/group levels); 2) Cubic SVM; 3) NN with 20 hidden layers; 4) DT with 100 splits; and 5) quadratic DA. Sparse representation had comparable accuracies to other classifiers (see **Appendix B**, **Table B1**), yet it was excluded from further analyses because of its low efficiency (i.e., >2 orders of magnitude greater that other classifiers for both training and testing), and was thus considered undesirable for use in real-time monitoring with the goal of adaptive activity detection. Naïve Bayes was also excluded because of consistently low accuracy (~60-75%) compared to the other algorithms (~80-99%).

Four different dataset generalization approaches, listed below, were used to evaluate the selected classifiers, with a variety of approaches used to assess the robustness of the classifiers. For each approach, results are presented for the testing datasets.

1. All SIB labels and non-SIB activity labels were balanced, such that the training:validating:testing ratio was 6:2:2.
2. The number of test points for each SIB type equaled the number of training points.
3. All remaining SIB observations that were excluded from training/validating were selected for testing, and the number of non-SIB activity samples equaled the maximum samples of SIB within a label.
4. Unbalanced, randomly selected data, such that the training:validating:testing ratio was 6:2:2, and the test data reflected the ratio of SIB:non-SIB during the entire session.

Results

Behaviors

Eleven participants, aged 5-14 years (M = 9.5, SD = 3.0, 3 females, 7 males) completed the study along with their caregiver(s). More than 1,000 minutes of data were obtained (session duration = 35-147 minutes) and >200 episodes of SIB were observed. Ten of the 11 participants exhibited

SIB (participants 1-4 and 6-11). Twenty-three different behaviors were exhibited, including 18 types of SIB and five types of stereotypical motor movements (SMM, summarized in **Table 2**). Note that the first participant had two different sensor configurations in one session; the child did not want to wear lower body sensors in initial data collection, yet the pocket and ankle sensors were placed on the child in time to capture all three SIB types. Also note that Participant 8 would only wear wrist and pocket sensors, and Participants 8 and 9 were siblings (each wore half of the sensors in consecutive sessions).

Data from participants with similar behaviors were grouped for analyses (**Table 3**). Groups 1-8 represent different behavior types (e.g., SIB for Participants 9 and 10, in G2, both involved hitting their heads against the wall). Variations of the groupings were included (e.g., G3 and G4, or G5-7), to evaluate the performance of classifiers on datasets that primarily represented a shared behavior (fewer participants included) and datasets that had additional forms of SIB along with the shared behavior (additional participants included). G9 contained data from participants who wore at least one sensor in each location (wrist, waist, pocket, and ankle). G10 included data from all participants, but only from the one sensor worn by all participants (the wrist).

Table 2. SIB displayed and sensors worn by each participant

Participant #	Behavior(s)*	Total Duration of Behavior (secs)	Sensors Worn
P1	Repeated foot to surface (1)	13	Wrist, waist (part 1) Wrist, waist, pockets, ankle (part 2)
	Repeated hand to surface (2)	6	
	Head hitting –with object (3)	20	
P2	Finger picking (4)	87	Wrist, waist, pockets, ankles
	Scratching (5)	28	
P3	Heel to surface (1)	66	Wrist, waist, pockets, ankles
	Hand to surface (2)	7	
P4	Self-biting (hands, arms) (9)	301	Wrist, waist, pockets, ankles
	Self-hitting (10)	80	
	Pulling teeth (11)	33	
	Eye-gouging (12)	79	
	Jabbing pelvic region (13)	16	

	Jabbing throat – location of prior tracheotomy (14)	46	
	Hitting chin/jaw with heel of hand (15)	66	
P6	Foot to surface (1)	2	Wrist, waist, pockets, ankles
	Hand to surface (2)	15	
	Repeatedly pulling on teeth using string/object (17)	256	
	Blowing on fingertips (16)	322	
	Spinning (18),	155	
	Flapping (19),	14	
	Jumping/flapping arms (20),	25	
	Jump/spin (21)	6	
P7	Finger picking (4)	322	Wrist, waist, pockets, ankles
P8	Foot to object (1)	2	Wrists, pockets
	Hand to surface (2)	4	
	Throwing body against object or surface (6)	22	
P9	Finger picking (4)	229	Wrist, waist, ankles
	Lip picking (7)	13	
	Head to wall (8)	14	
P10	Hands to surface (2)	9	Wrist, waist, pockets, ankles
	Finger Picking (4)	9	
	Scratching (5)	2	
	Head to wall (8)	20	
	Self-biting (9)	39	
	Self-hitting (10)	4	
	Eye-gauging (12)	58	
	Pulling ear (22)	209	
	Flapping (19)	5	
P11	Finger picking (4)	97	Wrist, waist, pockets,

	Hair pulling (23)	73	ankles

* Labels for each behavior are given in parentheses

Table 3. Groups of participants representing a shared behavior.

Group #	Shared Behavior(s)	Participants	Common Sensors Worn
G1	Foot to surface	1,3	Wrist, waist, pockets, ankle
G2	Head to wall	9,10	Wrist, waist, ankles
G3	Hand to surface	1, 3, 6, 8	Wrist, pockets
G4	Hand to surface	1, 3, 6, 8, 10	Wrist, pockets
G5	Finger picking	2, 7	Wrist, waist, pocket, ankles
G6	Finger picking	2, 7, 9, 11	Wrist, waist, ankles
G7	Finger picking	2, 7, 11	Wrist, waist, pocket, ankles
G8	Self-Biting	4,10	Wrist, waist, pocket, ankles
G9	All behaviors, measured with maximum sensors	2, 3, 4, 6, 7, 10, 11	Wrist, waist, pocket, ankles
G10	All behaviors from all participants	1 – 11	Wrist

SIB Detection (0,1 Classification)

At the individual level, two classifiers (kNN and SVM) consistently outperformed the other methods in terms of accuracy, specificity, and F-scores, while also achieving low offline classification times (>8.62 kHz, or ~0.1 msec/classification on average) when discriminating between SIB and non-SIB. Accuracy of kNN classifications on balanced test data (dataset generalization approach 1) ranged from 76.7 to 99.1%, with a mean of 93.0% (out of the number of individuals specified in Table II), and F-scores ranged from 0.75 to 0.99. Accuracy of the SVM classifier ranged from 74.3 to 98.5%, with a mean of 92.2% (out of the number of individuals specified in Table II), and F-scores from 0.75 to 0.99. **Table 4** summarizes performance metrics across participants and dataset generalization approaches. Results, in terms classifier performance, were generally consistent across all four dataset generalization approaches, though with a slight decrease evident when using unbalanced data (dataset generalization approach 4). All individually-trained classifiers, though, yielded high accuracy, specificity, and F-scores, even when testing all remaining SIB data (dataset generalization approach 3) and new unbalanced data with a low ratio of SIB:non-SIB events (dataset generalization approach 4).

When applied to participants grouped by similar behaviors, kNN and SVM methods again resulted in the highest classification accuracy, specificity, and F-scores. Of note, the participants were grouped according to similar behavior, though other types of SIB were still included and grouped as the same label, without differentiating between types of SIB. The groups represented the predominant SIB shared among participants, excluding specific labels beyond SIB detection. The kNN classifier had the highest accuracy for participants who demonstrated kicking (90.7%, G1), head banging (94.6%, G2), hitting (83.2%, G3), finger picking for the group with fewer participants (93.1%, G5), self-biting (78.7%, G8), participants wearing all sensors (82.6%, G9) and all participants (64.7%, G10). SVM was more accurate than kNN when classifying a larger number of participants who finger pick (93.1% for G6 and 94.3% for G7). The lowest accuracy was found for the group with participants who self-bite (G8, all classifiers <80% for balanced test data) and the group of all participants using one sensor (G10, all classifiers <65% for balanced test data). Though kNN and SVM consistently outperformed other classifiers, DT had the most efficient mean classification time at the group level (7.40×10^{-7} seconds per classification), which paralleled results at the individual level. Classifier performance was poorer at the group level when compared to individual levels. This trend was present overall, with the range of mean group accuracy levels = 71.9-91.5% compared to the range of individual accuracy levels (85.1-93.0%, **Table 4**). Classification with unbalanced data (dataset generalization approach 4) resulted in particularly poorer outcomes for groups of participants than for individual participants. For example, most group-level F-scores were <0.4.

Table 4. Mean classification performance across all participants for (0,1) labels.*****

Test	Method	Accuracy	Specificity	F-Score	PredTime (kHz)
1	kNN	0.930	0.960	0.927	8.62
2		0.928	0.950	0.925	46.3
3		0.925	0.946	0.926	78.1
4		0.925	0.946	0.926	75.2
1	SVM	0.922	0.911	0.923	34.5
2		0.919	0.909	0.920	69.0
3		0.918	0.906	0.922	123
4		0.918	0.906	0.922	116
1	NN	0.881	0.861	0.885	28.3
2		0.870	0.848	0.873	84.7
3		0.875	0.851	0.882	398
4		0.875	0.851	0.882	398
1	DT	0.882	0.885	0.882	13.5
2		0.889	0.894	0.888	140
3		0.885	0.885	0.889	565
4		0.885	0.885	0.889	459
1	DA	0.858	0.814	0.866	11.8
2		0.851	0.804	0.859	123
3		0.854	0.802	0.867	448
4		0.854	0.802	0.867	412

*Test = dataset generalization approach; Method = classification method; PredTime = time to complete classification on one observation offline.

SIB Types (Labels 0-23)

Use of kNN and SVM classifiers yielded superior performance in classifying the distinct SIB type (labeled 0-23), with accuracies at the individual level from 73.1-99.3% and 74.8-99.3% respectively. **Table 5** summarizes individual performance metrics (dataset generalization approach 1) using the SVM classifier, which had the highest accuracy; see **Appendix B, Table B2** for a summary of results using kNN. Across individuals, the highest mean accuracies were highest for hair pulling (97.2%) and head hitting (97.0%), though scratching had the highest specificity (99.6%) and F-score (0.996). The lowest mean accuracy for SVM (74.8%) was found when classifying pulling teeth with an object. Classifiers for SMM also performed with low mean accuracy (<75% for blowing on fingertips, spinning, jumping/flapping combination, and jumping/spinning combination).

For group-level classifiers, kNN and SVM remained the highest performing algorithms, with accuracies ranging from 73.7% (G3) to 97.0% (G7) for kNN, and from 75.3% (G3) to 95.9% (G7) for SVM, excluding G10. The latter group, which included all participants and one shared sensor (wrist), had substantially lower accuracy compared to other groups (48.8% for kNN, and 32.1% for SVM). Individually, G1, G2, G5, G6, and G7 had accuracies over 90% with balanced data, while G4, G8 and G9 had accuracies ~80%. When including as many participants as feasible with a maximum number of sensors (G9), kNN performed best, with 87.6% accuracy.

Table 6 summarizes performance metrics for the kNN classifier (dataset generalization approach 1), which had the highest accuracy for G10 (all participants). Behaviors that were relatively better classified included lip picking (92.4%) and scratching (89.6%), followed by ear pulling (83.1%). In spite of low accuracy evident for other behaviors (11.7 - 71.4%), specificity was high for all behaviors (>93%). Group-level results were also somewhat similar to those at the individual level, with high accuracy for classifying scratching and with consistently high specificity across all behaviors. In general, group-level classification yielded poorer accuracy and F-scores than at the individual level. There were also some common misclassified behaviors, specifically non-SIB events misclassified as finger picking or "hand to surface", along with "hand to surface" misclassified as throwing body on object/surface (see **Appendix**, **Table B3** for a complete confusion matrix).

Table 5. Mean SVM classification performance across all participants, for each behavior, (0-23) labels**.***

Behavior	Label	Accuracy	Specificity	F-score
Non-SIB events	0	0.920	0.974	0.877
Foot-to-surface	1	0.894	0.975	0.915
Fist-to-surface	2	0.901	0.979	0.916
Head hitting with hands/object	3	0.970	0.988	0.982
Picking fingers	4	0.957	0.969	0.964
Scratching	5	0.958	0.996	0.996
Throwing body on object/surface	6	0.917	0.956	0.916
Picking lip	7	0.957	0.990	0.962
Head banging (against wall)	8	0.945	0.985	0.947

Self-biting	9	0.898	0.983	0.901
Self-hitting	10	0.898	0.993	0.898
Pulling teeth (with hands)	11	0.862	0.982	0.925
Eye gauging	12	0.898	0.989	0.949
Jabbing pelvis	13	0.862	0.983	0.896
Jabbing throat	14	0.862	0.983	0.928
Knocking jaw	15	0.862	0.984	0.900
SMM - blowing on fingertips	16	0.748	0.972	0.825
Pulling teeth (with object)	17	0.748	0.962	0.811
SMM- spinning	18	0.748	0.962	0.685
SMM - flapping	19	0.841	0.974	0.823
SMM - jumping/flapping	20	0.748	0.955	0.638
SMM - jumping/spinning	21	0.748	0.983	0.743
Pulling ear	22	0.934	0.990	0.928
Pulling hair	23	0.972	0.992	0.989

*Reported performance using dataset generalization approach 1, unbalanced data

Table 6. kNN classification performance for each behavior, Group 10 (All Participants), (0-23) labels.*

Behavior	Label	Accuracy	Specificity	F-score
Non-SIB events	0	0.117	0.930	0.151
Foot to surface	1	0.330	0.953	0.404
Fist to surface	2	0.231	0.953	0.267
Head hitting with hands/object	3	0.553	0.976	0.603
Picking fingers	4	0.520	0.972	0.588
Scratching	5	0.896	0.995	0.899
Throwing body on object/surface	6	0.354	0.966	0.385
Picking lip	7	0.924	0.997	0.921
Head banging (against wall)	8	0.554	0.975	0.615
Self-biting	9	0.343	0.969	0.355
Self-hitting	10	0.226	0.959	0.247
Pulling teeth (with hands)	11	0.640	0.983	0.654
Eye gauging	12	0.582	0.980	0.609
Jabbing pelvis	13	0.677	0.986	0.677
Jabbing throat	14	0.772	0.991	0.713
Knocking jaw	15	0.636	0.984	0.626
SMM - blowing on fingertips	16	0.274	0.974	0.243
Pulling teeth (with object)	17	0.525	0.984	0.448
SMM- spinning	18	0.375	0.983	0.282
SMM - flapping	19	0.398	0.982	0.317
SMM - jumping/flapping	20	0.444	0.985	0.336
SMM - jumping/spinning	21	0.550	0.997	0.239
Pulling ear	22	0.831	0.996	0.545
Pulling hair	23	0.714	0.993	0.524

*Reported performance using dataset generalization approach 1, unbalanced data

Discussion

Several machine learning algorithms detected SIB and classified diverse behavior types with high accuracy and efficiency, both for individuals and groups of children with ASD. Our methods were trained with a more numerous sample (10 participants) than in prior work on the classification of ASD behavior (Coronato et al., 2014; Goodwin et al., 2014; Goodwin et al., 2011; Min & Tewfik, 2010). Our sample was still relatively small, though it was consistent with the exploratory nature of the work and it included actual SIB (versus simulated) data. Two particular algorithms, kNN and SVM, were consistent in classifying behaviors with superior performance. This finding was evident for individual and group-level classifiers, and for both limited labels (0,1) and all labeled behaviors (0-23). The resulting levels of accuracy (up to 99.1% for individuals and up to 97.0% for groups) either match (Coronato et al., 2014; Goodwin et al., 2008; Min et al., 2009) or surpass (Goncalves et al., 2012; Plötz et al., 2012) those in earlier reports addressing other types of movement in ASD. There was a drop in accuracy from individual level accuracies to group level accuracies for specific behaviors (e.g., from ~95-97% to ~55% for head banging/hitting), which is supported by prior work that found poorer performance outcomes in group-level classifiers (Ozdenizci et al., 2018). Individual participants often showed different presentations of the same type of SIB (e.g., head-banging the back of the head versus the side or front), which may explain the higher individual-specific behavioral classifiers. We also found that SVMs outperformed the accuracy of DTs, by up to 11% when comparing means of the (0,1) label scheme. This result is consistent with a previous finding that SVM surpassed the performance of DTs by ~6-10% when classifying SMM (Goodwin et al., 2014). Consistently high specificity was also found (i.e., >.80 for SIB detection and classification), suggesting the potential for a low false-alarm rate for classifiers implemented into a real-time monitoring system. Both the kNN and SVM classifiers are thus promising for use in behavioral monitoring, with the kNN offering simple implementation and SVMs offering a robust classification method (Wu et al., 2008).

Despite the high level of accuracy obtained with kNN and SVMs in general, performance outcomes differed between specific SIB behaviors. At the group level, movements with larger magnitudes (e.g., hand/foot to surface, self-hitting) were detected more poorly, with lower accuracies and F-scores, than finer movements (e.g., pulling hair, picking). Throwing the body onto an object/surface, for example, was commonly misclassified as “hand to surface” behavior in G10 (though not at individual levels). This misclassification could have been because one participant’s movement when hitting a table resembled the movement from another participant throwing himself on that table (i.e., the upper body moves towards the table first for both types of SIB). Contrasting this finding on low accuracy with larger movements, scratching was accurately classified even when grouped with other behaviors in G10. Performance for this SIB was better here (F = 0.90) than in prior work using wrist accelerometers to detect a similar motion, nocturnal scratching (F = 0.68; Moreau et al., 2018). Though our results suggest the feasibility of monitoring motor movements as subtle as scratching, which are often difficult to notice through observation, enhancing detection accuracy for larger motions requires further exploration.

Previous research indicated that the presence of SMMs may be predictive of the occurrence of SIB (Minshawi et al., 2014). Though SMMs were not of particular interest here, they were displayed by both P6 and P10, and were therefore labeled as behaviors potentially indicative of SIB onset. However, SMMs were detected here with lower accuracy than SIB in the group level classifiers, commonly misclassified as other SMM or non-SIB. For example, in G10, blowing on fingertips repeatedly (an SMM) was commonly misclassified as non-SIB and spinning was most commonly misclassified as blowing on fingertips, likely because this SMM was a subtle motion that often preceded spinning. The SMM of flapping was also commonly misclassified as blowing on fingers, possibly due to the similar motion of drawing the arm towards the face or similar rhythmic tendencies, as well as flapping/jumping due to overlapping motions. Prior research on SMM detection (Goncalves et al., 2012) should thus not be assumed to generalize to SIB. Though these noted misclassifications may have led to low F-scores for SMMs, similar patterns of behavior could be used in future work to identify sequences of SMMs and SIB. When an SMM was exhibited by more than one participant (two, for flapping), accuracy increased at the group level (G10), which indicates that both SIB and SMM could be differentiated in automatic monitoring if trained with multiple examples. Thus, future work may require data from both SMMs and SIB to accurately distinguish between these behaviors.

Performance outcomes for different behaviors may be contingent on the extraneous activity of the child being monitored. For example, participants who exhibited SIBs that were detected accurately (e.g., picking or scratching, as described above) were typically sitting in one place long enough to begin and continue the SIB. Having fewer motions surrounding the SIB, with little "noise" from unrelated activity, likely supported this high accuracy. Other participants tended to be more active, moving from SIB (e.g., self-hitting) into another similar motion (e.g., dancing). Behaviors shown by more active participants (e.g., P1, P3, P4, P6, P8, P10) were indeed commonly misclassified, in contrast to the high accuracy obtained for subtle SIB by less active participants (e.g., P2, P7, P9, P11). Several caregivers of participants in the former group also alluded to the presence of attention-deficit/hyperactivity disorder in their child, which could have contributed to the high levels of activity. Behaviors with the lowest accuracies (e.g., pulling teeth with an object, most SMMs) were shown exclusively by a highly active child (P6). These behaviors may have been poorly classified because they were exhibited by only the one individual, and this individual was consistently moving around the room. These outcomes support the importance of including a range of activity levels when training and testing SIB classification models.

Our results also highlighted the importance of sensor configurations for individual level models, as shown by the differences in classifier performance for two sensor configurations of P1. Poorer performance was obtained with data from the first part of the session, when P1 was wearing only the wrist and waist sensors, versus the second part of P1 session, with additional sensors in the pockets and on the ankles. For example, when detecting SIB vs. non-SIB, kNN yielded a classification accuracy = 94.0% and F = 0.94 when the child was wearing two sensors (wrist and waist), which increased to 99.1% and F = 0.99 when wearing additional pocket and ankle sensors. The higher accuracy in this second part of the session is likely due to the added sensors; notably, the latter reflected the configuration suggested in prior work (Plötz et al., 2012) for optimal activity classification (i.e., wrist and ankle sensors). The added sensors were on the lower body, where the child showed SIB. Though the wrist and waist sensors provided sufficient data for classifying head banging in the first part of the P1 session, the ankle sensors provided necessary data for lower body SIB. Additional data for specific SIB types should be collected to determine optimal placement according to behavior.

Sensor configuration also appeared important based on group-level results. For example, G9 and G10 differed both in sensor configuration and in the number of participants, with G10 having fewer sensors (1 vs. 5) yet three additional participants. Accuracy in G10 was 48.8%, substantially lower than the 87.6% for G9. Individually, however, the participants added from G9 to G10 each had mean accuracies of >91.7% (labels 0-23). Thus, the substantial drop in accuracy was likely a result of the loss of five sensors (or 15 channels of data) versus the addition of three participants. More than one sensor may therefore be necessary when classifying groups of diverse participants.

In addition to accuracy, efficiency is critical when working toward a real-time monitoring system. Efficiency has multiple important components in the current context, related to data processing and model development (training). Regarding data processing, we used featureless data, involving minimal processing, and with filtered sensor channels directly input into these classifiers as in Bastani, Kim, et al. (2016). Featureless methods were also considered preferable given the lack of prior work to guide appropriate SIB feature extraction. With individual-level accuracies up to 99.1%, the results of our work show the potential of featureless methods for efficient monitoring in continuous online analyses. Efficiency in the training stage was addressed through examining the minimum number of training points required for accurate results. Our findings suggest that only 400-500 data points (i.e., <10 seconds, cumulatively, of SIB data acquired at 60 Hz) for each SIB type were needed to build individual-level models, or models with participants grouped according to primary behavior types. These training data yielded classifiers that could accurately distinguish SIB within a substantially longer testing set, as in dataset generalization approach 3. This finding has implications for analysts who typically conduct FAs, as it could support a more efficient data collection process.

Along with the quantitative results described above, qualitative results were provided through participant feedback. Though most participants were minimally verbal, two children did provide feedback on their experiences with accelerometers. For example, one 8-year-old participant stated that he did not like the shoe sensors, “since [he doesn’t] wear [his] shoes all the time” and would “have to take them off”. This same participant noted that the one on his back fell off when he played. He said he “didn’t notice the pockets [sensors] at all”. Another participant stated that he did not like them on his shoes, and he instead preferred them in his pocket or on his wrists. Participating children also frequently expressed interest in one material over another, even if the participant was minimally verbal or nonverbal. Children most commonly selected fleece and sequins. (Note that we created a sleeve out of the fleece, such that only fleece made contact with the skin, and we wrapped the sequin wristband, with soft material underneath, around the wrist sensor.) Several children also showed a preference for the sensors decorated with his/her favorite stickers, especially if the child decorated the sensor. Several children removed or threw the sensors if emotions escalated (e.g., if the caregiver initiated SOAP). However, we observed that the children commonly ignored pocket sensors or the waist sensor on the lower back, and that these sensors were less likely to be removed or targeted during aggression. Though not intentionally removed, the lower back/waist sensor tended to fall off with clothing adjustment. These findings could help guide future work on wearable technology among children with ASD. Overall, both explicit and observed sources of feedback suggested that pocket sensors should be considered as a feasible method to movement monitoring.

Limitations and future directions

This study offers initial results in support of developing an SIB monitoring system. However, further work is recommended to obtain a group-level classifier with stronger performance than was found here (individual-level classifiers were highly accurate, yet may be less efficient for practical application with required training for each individual). The desired group-level performance may be achieved with variations of the evaluated classifiers. SRC was not explored beyond validation due to its low efficiency. However, SRC may provide superior performance if paired with novel sparse approximation methods, such as the “greedy Bayesian” approach used in Bastani et al. (2016). One possible avenue to improve group-level accuracy is through additional processing techniques, including feature extraction and selection. Featureless methods were used here for efficiency (as noted early) and initial model exploration, yet feature extraction and selection could improve other performance metrics. Classification based on features has been found in prior work on SMM detection to yield better performance than featureless methods (Rad et al., 2016). Models including SIB characteristics (as example “features”) might provide further information about SIB and thereby contribute towards enhanced accuracy. Further, accuracy may be limited by the rate at which researchers viewed and labeled the video footage (1 frame/second), though trained observers in traditional tracking methods (i.e., functional assessments) may track behaviors in even larger intervals (i.e., tally mark estimates for each hour of the day). Resulting annotations are thus of finer resolution than the typical accepted standard of functional assessments, yet accuracy may improve if viewing the video data in higher resolution.

This work represents a preliminary examination of SIB classification with machine learning algorithms. As an initial effort towards SIB monitoring, the participants were monitored within a specific and controlled setting (i.e., with limited space for activity and interaction with others). Though a necessary first step to capture "ground truth" data (via video) in a safe setting, movement data were collected only at a child study center. As such, our results may not generalize to other environments, specifically in areas with high level of activity (e.g., playgrounds, or gym class in school) that were not represented in the child study center. To create an SIB monitoring system that transfers across contexts, future work should include data collected in diverse locations where the individual typically shows SIB (e.g., at school or in the community).

The present study is the first known application of diverse methods to efficiently and accurately classify SIB. However, these methods should be extended to further consider real-world applications. Future work could include additional modeling techniques, such as sliding windows. The current models yielded offline classification rates that are faster than the sampling rate (60 Hz), typically exceeding 8000 decisions (classifications) per second. Though this finding is indicative of highly efficient classification, it could lead to overwhelming caregivers with unnecessary information. Instead, it may be more effective to provide a single decision (i.e., label = SIB or non-SIB) within a set time frame. For example, a decision could be provided every two seconds, compiling across the more frequent decisions made in that window. With this technique, caregivers could be alerted of an SIB event in a meaningful timeframe.

Conclusions

Our results are the first, to our knowledge, that indicate the feasibility of detecting a range of SIB with high accuracy, specificity, and efficiency, using wearable sensors (accelerometers). At the individual level, classification models based on kNN and SVM detected SIB with accuracy up to 99.1% (mean accuracy across individuals of ~93%). These same classifiers were effective at the group level, for different behavioral types, with mean accuracy up to 97%. Results of this study support the potential for developing an effective monitoring system for SIB, one the most dangerous concerns in ASD.

Acknowledgments

The author was supported by a National Science Foundation Graduate Research Fellowship. However, this agency did not have any involvement in data analysis, interpretation, or the decision to publish.

References

Albinali, F., Goodwin, M. S., & Intille, S. S. (2009). *Recognizing stereotypical motor movements in the laboratory and classroom: A case study with children on the autism spectrum.* Paper presented at the Proceedings of the 11th international conference on Ubiquitous computing.

Allen, K. D., & Warzak, W. J. (2000). The problem of parental nonadherence in clinical behavior analysis: Effective treatment is not enough. *Journal of Applied Behavior Analysis, 33*(3), 373-391. doi:10.1901/jaba.2000.33-373

Amaral, D. G., Schumann, C. M., & Nordahl, C. W. (2008). Neuroanatomy of autism. *Trends in Neurosciences, 31*(3), 137-145. doi:https://doi.org/10.1016/j.tins.2007.12.005

Banaee, H., Ahmed, M., & Loutfi, A. J. S. (2013). Data mining for wearable sensors in health monitoring systems: A review of recent trends and challenges. *13*(12), 17472-17500.

Bastani, K., Kim, S., Kong, Z. J., Nussbaum, M. A., & Huang, W. (2016). Online classification and sensor selection optimization with applications to human material handling tasks using wearable sensing technologies. *IEEE Transactions on Human-Machine Systems, 46*(4), 485-497. doi:10.1109/THMS.2016.2537747

Bastani, K., Rao, P. K., & Kong, Z. (2016). An online sparse estimation-based classification approach for real-time monitoring in advanced manufacturing processes from heterogeneous sensor data. *IIE Transactions, 48*(7), 579. doi:10.1080/0740817X.2015.1122254

Bellini, S., & Akullian, J. (2007). A meta-analysis of video modeling and video self-modeling interventions for children and adolescents with autism spectrum disorders. *Exceptional Children, 73*(3), 264-287.

Bone, D., Goodwin, M. S., Black, M. P., Lee, C.-C., Audhkhasi, K., & Narayanan, S. (2015). Applying machine learning to facilitate autism diagnostics: Pitfalls and promises. *Journal of Autism and Developmental Disorders, 45*(5), 1121-1136. doi:10.1007/s10803-014-2268-6

Bulling, A., Blanke, U., & Schiele, B. (2014). A tutorial on human activity recognition using body-worn inertial sensors. *ACM Computing Surveys (CSUR), 46*(3), 33.

Cabibihan, J.-J., Javed, H., Aldosari, M., Frazier, T., & Elbashir, H. (2017). Sensing technologies for autism spectrum disorder screening and intervention. *Sensors, 17*(1), 46.

Cantin-Garside, K., Valdez, R. S., Nussbaum, M. A., White, S., Kim, S., Kim, C. D., & Fortes, D. M. (2018). *Exploring challenges of monitoring technology and self-injurious behavior in autism spectrum disorder.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Chen, Y.-H., Rodgers, J., & McConachie, H. (2009). Restricted and repetitive behaviours, sensory processing and cognitive style in children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39*(4), 635-642. doi:10.1007/s10803-008-0663-6

Coronato, A., De Pietro, G., & Paragliola, G. (2014). A situation-aware system for the detection of motion disorders of patients with autism spectrum disorders. *Expert Systems with Applications, 41*(17), 7868-7877.

Corporation, P. (1999). Wechsler abbreviated scale of intelligence. In: Author San Antonio, TX.

Dracobly, J. D., Dozier, C. L., Briggs, A. M., & Juanico, J. F. (2018). Reliability and validity of indirect assessment outcomes: Experts versus caregivers. *Learning and Motivation, 62*, 77-90. doi:https://doi.org/10.1016/j.lmot.2017.02.007

Dunlap, G., Newton, J. S., Fox, L., Benito, N., & Vaughn, B. (2001). Family involvement in functional assessment and positive behavior support. *Focus on Autism and Other Developmental Disabilities, 16*(4), 215-221.

Gaonkar, B., & Davatzikos, C. (2013). Analytic estimation of statistical significance maps for support vector machine based multi-variate image analysis and classification. *NeuroImage, 78*, 270-283. doi:https://doi.org/10.1016/j.neuroimage.2013.03.066

Goncalves, N., Rodrigues, J. L., Costa, S., & Soares, F. (2012). *Preliminary study on determining stereotypical motor movements*, United States.

Goncalves, N., Rodrigues, J. L., Costa, S., & Soares, F. (2012, 2012). *Automatic detection of stereotyped hand flapping movements: Two different approaches*.

Goodwin, M., Haghighi, M., Tang, Q., Akcakaya, M., Erdogmus, D., & Intille, S. (2014, 2014). *Moving towards a real-time system for automatically recognizing stereotypical motor movements in individuals on the autism spectrum using wireless accelerometry*.

Goodwin, M. S., Intille, S. S., Albinali, F., & Velicer, W. F. (2011). Automated detection of stereotypical motor movements. *Journal of Autism and Developmental Disorders, 41*(6), 770-782. doi:10.1007/s10803-010-1102-z

Goodwin, M. S., Intille, S. S., Velicer, W. F., & Groden, J. (2008). *Sensor-enabled detection of stereotypical motor movements in persons with autism spectrum disorder.* Paper presented at the Proceedings of the 7th international conference on Interaction design and children.

Gowen, E., & Hamilton, A. (2013). Motor abilities in autism: A review using a computational context. *Journal of Autism and Developmental Disorders, 43*(2), 323. doi:10.1007/sl0803-012-1574-0

Großekathöfer, U., Manyakov, N. V., Mihajlović, V., Pandina, G., Skalkin, A., Ness, S., . . . Goodwin, M. S. (2017). Automated detection of stereotypical motor movements in autism spectrum disorder using recurrence quantification analysis. *Frontiers in Neuroinformatics, 11*, 9. doi:10.3389/fninf.2017.00009

Hill, A. P., Zuckerman, K. E., Hagen, A. D., Kriz, D. J., Duvall, S. W., van Santen, J., . . . Fombonne, E. (2014). Aggressive behavior problems in children with autism spectrum disorders: Prevalence and correlates in a large clinical sample. *Research in Autism Spectrum Disorders, 8*(9), 1121-1133. doi:10.1016/j.rasd.2014.05.006

Iwata, B. A., Pace, G. M., Dorsey, M. F., Zarcone, J. R., Vollmer, T. R., Smith, R. G., . . . Willis, K. D. (1994). The functions of self-injurious behavior: An experimental-epidemiological analysis. *27*(2), 215-240. doi:doi:10.1901/jaba.1994.27-215

Johnson, C. R., Butter, E. M., Handen, B. L., Sukhodolsky, D. G., Mulick, J., Lecavalier, L., . . . McDougle, C. J. (2009). Standardised observation analogue procedure (soap) for assessing parent and child behaviours in clinical trials. *J Intellect Dev Disabil, 34*(3), 230-238. doi:10.1080/13668250903074471

Kalb, L. G., Vasa, R. A., Ballard, E. D., Woods, S., Goldstein, M., & Wilcox, H. C. (2016). Epidemiology of injury-related emergency department visits in the us among youth with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 46*(8), 2756-2763. doi:10.1007/s10803-016-2820-7

Kanne, S. M., & Mazurek, M. O. (2011). Aggression in children and adolescents with asd: Prevalence and risk factors. *Journal of Autism and Developmental Disorders, 41*(7), 926-937. doi:10.1007/s10803-010-1118-4

Kim, S., & Nussbaum, M. A. (2014). An evaluation of classification algorithms for manual material handling tasks based on data obtained using wearable technologies. *Ergonomics, 57*(7), 1040-1051. doi:10.1080/00140139.2014.907450

Kirby, A. V., Boyd, B. A., Williams, K. L., Faldowski, R. A., & Baranek, G. T. (2016). Sensory and repetitive behaviors among children with autism spectrum disorder at home. *Autism*, 1-13.

Kurtz, P. F., Chin, M. D., Huete, J. M., Tarbox, R. S., O'Connor, J. T., Paclawskyj, T. R., & Rush, K. S. (2003). Functional analysis and treatment of self‐injurious behavior in young children: A summary of 30 cases. *Journal of Applied Behavior Analysis, 36*(2), 205-219.

Lam, K. S., & Aman, M. G. (2007). The repetitive behavior scale-revised: Independent validation in individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 37*(5), 855-866.

Lord, C., Rutter, M., DiLavore, P., Risi, S., Gotham, K., & Bishop, S. (2012). Autism diagnostic observation schedule–2nd edition (ados-2). *Los Angeles, CA: Western Psychological Corporation*.

Marcu, G., Tassini, K., Carlson, Q., Goodwyn, J., Rivkin, G., Schaefer, K. J., . . . Kiesler, S. (2013). *Why do they still use paper?: Understanding data collection and use in autism education*. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Paris, France.

McLeod, A., Bochniewicz, E. M., Lum, P. S., Holley, R. J., Emmer, G., & Dromerick, A. W. (2016). Using wearable sensors and machine learning models to separate functional upper extremity use from walking-associated arm movements. *Archives of Physical Medicine and Rehabilitation, 97*(2), 224.

Miller, J. D., Beazer, M. S., & Hahn, M. E. (2013). Myoelectric walking mode classification for transtibial amputees. *IEEE Transactions on Biomedical Engineering, 60*(10), 2745-2750. doi:10.1109/TBME.2013.2264466

Min, C.-H. (2017). *Automatic detection and labeling of self-stimulatory behavioral patterns in children with autism spectrum disorder.* Paper presented at the Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE.

Min, C.-H., & Tewfik, A. H. (2010). *Automatic characterization and detection of behavioral patterns using linear predictive coding of accelerometer sensor data.* Paper presented at the Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE.

Min, C.-H., Tewfik, A. H., Kim, Y., & Menard, R. (2009, 2009). *Optimal sensor location for body sensor network to detect self-stimulatory behaviors of children with autism spectrum disorder*, United States.

Minshawi, N. F., Hurwitz, S., Fodstad, J. C., Biebl, S., Morriss, D. H., & McDougle, C. J. (2014). The association between self-injurious behaviors and autism spectrum disorders. *Psychology research and behavior management, 7*, 125-136. doi:10.2147/PRBM.S44635

Mittek, M. M., Carlson, J. D., Mora-Becerra, F., Psota, E. T., & Perez, L. D. (2015). In-home behavioral monitoring using simultaneous localization and activity detection. *Biomedical sciences instrumentation, 51*, 289.

Moreau, A., Anderer, P., Ross, M., Cerny, A., Almazan, T. H., & Peterson, B. (2018). Detection of nocturnal scratching movements in patients with atopic dermatitis using accelerometers and recurrent neural networks. *IEEE Journal of Biomedical and Health Informatics, 22*(4), 1011-1018. doi:10.1109/JBHI.2017.2710798

Ozdenizci, O., Cumpanasoiu, C., Mazefsky, C., Siegel, M., Erdogmus, D., Ioannidis, S., & Goodwin, M. S. (2018). Time-series prediction of proximal aggression onset in minimally-verbal youth with autism spectrum disorder using physiological biosignals. *arXiv preprint arXiv:1809.09948*.

Pace, G. M., Iwata, B. A., Edwards, G. L., & McCosh, K. C. (1986). Stimulus fading and transfer in the treatment of self‐restraint and self‐injurious behavior. *Journal of Applied Behavior AnalysiS, 19*(4), 381-389.

Pelios, L., Morren, J., Tesch, D., & Axelrod, S. (1999). The impact of functional analysis methodology on treatment choice for self‐injurious and aggressive behavior. *Journal of Applied Behavior Analysis, 32*(2), 185-195.

Plötz, T., Hammerla, N. Y., & Olivier, P. (2011). *Feature learning for activity recognition in ubiquitous computing.* Paper presented at the IJCAI Proceedings-International Joint Conference on Artificial Intelligence.

Plötz, T., Hammerla, N. Y., Rozga, A., Reavis, A., Call, N., & Abowd, G. D. (2012). *Automatic assessment of problem behavior in individuals with developmental disabilities.* Paper presented at the Proceedings of the 2012 ACM Conference on Ubiquitous Computing.

Poliker, R. (2006). Pattern recognition. In *Wiley Encyclopedia of Biomedical Engineering*: John Wiley & Sons, Inc.

Preece, S. J., Goulermas, J. Y., Kenney, L. P., Howard, D., Meijer, K., & Crompton, R. (2009). Activity identification using body-mounted sensors—a review of classification techniques. *Physiological Measurement, 30*(4), R1.

Pugliese, C. E., Kenworthy, L., Bal, V. H., Wallace, G. L., Yerys, B. E., Maddox, B. B., . . . Anthony, L. G. (2015). Replication and comparison of the newly proposed ados-2, module 4 algorithm in asd without id: A multi-site study. *Journal of Autism and Developmental Disorders, 45*(12), 3919-3931. doi:10.1007/s10803-015-2586-3

Rad, N. M., Bizzego, A., Kia, S. M., Jurman, G., Venuti, P., & Furlanello, C. (2015). Convolutional neural network for stereotypical motor movement detection in autism. *arXiv preprint arXiv:1511.01865*.

Rad, N. M., Furlanello, C., & Kessler, F. B. (2016). Applying deep learning to stereotypical motor movement detection in autism spectrum disorders.

Richards, C., Moss, J., Nelson, L., & Oliver, C. (2016). Persistence of self-injurious behaviour in autism spectrum disorder over 3 years: A prospective cohort study of risk markers. *J Neurodev Disord, 8*, 21. doi:10.1186/s11689-016-9153-x

Rojahn, J., Matson, J. L., Lott, D., Esbensen, A. J., & Smalls, Y. (2001). The behavior problems inventory: An instrument for the assessment of self-injury, stereotyped behavior, and aggression/destruction in individuals with developmental disabilities. *Journal of Autism and Developmental Disorders, 31*(6), 577-588.

Rooker, G. W., Hausman, N. L., Fisher, A. B., Gregory, M. K., Lawell, J. L., & Hagopian, L. P. (2018). Classification of injuries observed in functional classes of self-injurious behaviour. *62*(12), 1086-1096. doi:doi:10.1111/jir.12535

Schaeffer, K. M., Hamilton, K. A., & Johnson, W. L. B. (2016). Video self-modeling interventions for students with autism spectrum disorder. *Intervention in School and Clinic*, 1053451216630281.

Soares, D. A., Vannest, K. J., & Harrison, J. (2009). Computer aided self-monitoring to increase academic production and reduce self-injurious behavior in a child with autism. *Behavioral Interventions, 24*(3), 171-183. doi:10.1002/bin.283

Tarbox, J., Wilke, A. E., Najdowski, A. C., Findel-Pyles, R. S., Balasanyan, S., Caveney, A. C., . . . Tia, B. (2009). Comparing indirect, descriptive, and experimental functional assessments of challenging behavior in children with autism. *Journal of Developmental and Physical Disabilities, 21*(6), 493-514. doi:10.1007/s10882-009-9154-8

Taylor, L., Oliver, C., & Murphy, G. (2011). The chronicity of self-injurious behaviour: A long-term follow-up of a total population study. *Journal of Applied Research in Intellectual Disabilities, 24*(2), 105-117. doi:doi:10.1111/j.1468-3148.2010.00579.x

Trost, S. G., Zheng, Y., & Wong, W.-K. J. P. m. (2014). Machine learning for activity recognition: Hip versus wrist data. *35*(11), 2183.

Williams, S. K., Johnson, C., & Sukhodolsky, D. G. (2005). The role of the school psychologist in the inclusive education of school-age children with autism spectrum disorders. *Journal of School Psychology, 43*(2), 117-136.

Wright, J., Yang, A. Y., Ganesh, A., Sastry, S. S., Ma, Y. J. I. t. o. p. a., & intelligence, m. (2009). Robust face recognition via sparse representation. *31*(2), 210-227.

Wu, X., Kumar, V., Ross Quinlan, J., Ghosh, J., Yang, Q., Motoda, H., . . . Steinberg, D. (2008). Top 10 algorithms in data mining. *Knowledge and Information Systems, 14*(1), 1-37. doi:10.1007/s10115-007-0114-2

Zheng, Y.-L., Ding, X.-R., Poon, C. C. Y., Lo, B. P. L., Zhang, H., Zhou, X.-L., . . . Zhang, Y.-T. (2014). Unobtrusive sensing and wearable devices for health informatics. *IEEE Transactions on Biomedical Engineering, 61*(5), 1538-1554.

Chapter 4: Multi-Level Modeling with Nonlinear Metrics of Motor Variability to Identify Self-Injurious Behavior in Autism Spectrum Disorder

Abstract

Self-injurious behavior (SIB) is among the most dangerous concerns in autism spectrum disorder (ASD), often requiring detailed and tedious management methods. Smart monitoring could address the limitations of common management methods, specifically through automatically tracking behavioral patterns and identifying SIB, though the complex problem of classifying variable behavior should first be addressed. The current study aimed to address this need through developing and testing a group-level model that accounts for individual variability, as well as potential nonlinear trends in SIB movements. Ten participants with ASD and SIB engaged in free play and everyday tasks at a child study center. Body-worn accelerometers were secured on participants, and movement data from >200 episodes and 18 different types of SIB were collected. Lasso and principal component analysis were used to select features and reduce dimensions, respectively, which resulted in >65% explained variance and 7 of 12 selected principal components loading on nonlinear metrics. A multi-level logistic regression model classified SIB with >75% accuracy. Our findings imply that features consisting of nonlinear metrics of SIB movements significantly predict SIB. The model was superior to other commonly used classifiers, and classification performance was superior to results in prior reports. This work provides an approach to generating an accurate and interpretable group-level model for SIB identification, and supports the feasibility of developing a real-time SIB monitoring system that applies to a wide range of behaviors and individuals with ASD.

Keywords: *autism spectrum disorder, self-injurious behavior, principal components analysis, logistic regression, nonlinear motor variability*

Introduction

Autism spectrum disorder (ASD) is a pervasive neurodevelopmental disability marked by communicative, social and behavioral impairments (APA, 2013). The estimated prevalence of ASD increased from 1 in 150 in 2000 to 1 in 59 youth in 2014 reports (Baio et al., 2018; Rice, 2007). This increase in diagnoses has a concomitant increased need for early intervention and long-term care (Dawson et al., 2010; Lucyshyn et al., 2007), particularly for those without access to in-person options (Krahn, Klein Walker, & Correa-De-Araujo, 2015; Liptak et al., 2008; Vohra, Madhavan, Sambamoorthi, & St Peter, 2013). These needs are especially critical to address the severe concerns associated with ASD, especially self-injurious behavior (SIB), which includes head banging and self-hitting (Minshawi et al., 2014), and is reported in roughly half of people with ASD (McTiernan, Leader, Healy, & Mannion, 2011; Murphy, Healy, & Leader, 2009; Richards, Oliver, Nelson, & Moss, 2012). SIB is a leading cause of hospitalization among people with ASD, and can lead to physical damage such as lacerations and contusions (Minshawi et al., 2014; Rooker et al., 2018). These behaviors can be repetitive or rhythmic, though behavior presentations vary widely (Minshawi et al., 2014). Functions of the behavior (e.g., escaping a demand) also differ both between individuals and within an individual (Iwata et al., 1994; Minshawi, Hurwitz, Morriss, & McDougle, 2015). Applied behavioral analysis thus suggests that caregivers perform a functional assessment (FA) to determine possible triggers of SIB (Goodwin et al., 2014; Kirby, Boyd, Williams, Faldowski, & Baranek, 2016; Williams, Johnson, & Sukhodolsky, 2005).

As prescribed by applied behavioral analysis, FA is a critical first step towards creating an effective management plan for SIB (Goodwin et al., 2014; Kirby et al., 2016; Williams et al., 2005). To complete an FA, clinicians or trained caregivers observe and record details about events preceding, during, and following SIB (Williams et al., 2005). Clinicians can then use this information to make inferences about associated triggers to redirect, replace, or extinguish the behavior (Iwata et al., 1994; Pelios, Morren, Tesch, & Axelrod, 1999; Williams et al., 2005). However, FA results may have poor accuracy and validity. Accuracy may suffer if caregivers are not adequately trained, and if events are recalled after they occurred (Marcu et al., 2013; Tarbox et al., 2009). Observations can differ between caregivers and clinicians, and can be challenging to track consistently due to other stressors and contextual influences on behavior (Allen & Warzak, 2000; Dracobly, Dozier, Briggs, & Juanico, 2018; Dunlap, Newton, Fox, Benito, & Vaughn, 2001; Kirby et al., 2016; Marcu et al., 2013). Further, traditional manual methods are often inefficient (Dracobly et al., 2018; Marcu et al., 2013), which do not support the widespread need for care across contexts. SIB can evolve and often persists as the child ages (Kurtz et al., 2003), which requires continual support and vigilance, even after FA is completed. Management plans may need to adapt to functional or behavioral changes.

To address these challenges, an accurate and valid tracking system could be effective both for FA and to inform and evaluate management (see Chapter 3 for additional details). As shown by previous research in behavioral monitoring for other behaviors in ASD (Cabibihan, Javed, Aldosari, Frazier, & Elbashir, 2017; Goodwin, Intille, Albinali, & Velicer, 2011; Plötz et al., 2012), sensing technology could comprehensively, objectively, and accurately track movement for people with SIB. However, technology implementation in management has been limited to invasive and tedious video observation (Bellini & Akullian, 2007; Kirby et al., 2016; Schaeffer, Hamilton, & Johnson, 2016; Soares, Vannest, & Harrison, 2009). Nonwearable and wearable innovations, such as embedded camera systems or accelerometers in everyday items (e.g., cellphones), could continuously record data for SIB monitoring without requiring high levels of caregiver or clinician compliance (Cabibihan et al., 2017; Zheng et al., 2014). Nonwearable systems may address sensory sensitivity often existing among children with ASD and restrictive or repetitive movements (Chen, Rodgers, & McConachie, 2009), yet several challenges can arise. As examples, multiple devices may need to be installed to capture a sufficient field of view, and embedded technology could present a privacy concern if capturing caregiver movement (Goncalves, Rodrigues, Costa, & Soares, 2012). Wearable accelerometers address these limitations, and were selected for the current study to reflect caregiver preferences from our previous work (Cantin-Garside et al., 2018). Caregivers in that work indicated a need for collection methods applicable in school and at home, and suggested that children with SIB would accept wearable technology, assuming noninvasive, comfortable, and discrete attachment methods were possible (Cantin-Garside et al., 2018). Accelerometers have also been shown to provide sufficient data to detect repetitive motions among individuals with ASD, with 80-97% accuracy using wrist and/or back sensors (Goodwin, Intille, Velicer, & Groden, 2008; Min, Tewfik, Kim, & Menard, 2009), though use for SIB detection should be further explored.

In conjunction with the described wearable technology, “smart” SIB monitoring requires effective modeling. Earlier findings support the feasibility of tracking behaviors in ASD, specifically stereotypical motor movements (SMM) such as hand-flapping or rocking, which may relate to SIB and be similarly repetitive and rhythmic (Minshawi et al., 2014). Machine learning algorithms applied to accelerometry data – including decision trees (Albinali, Goodwin, & Intille, 2011), neural networks, (Coronato, De Pietro, & Paragliola, 2014; Rad, Furlanello, & Kessler, 2016), and support vector machines (Goodwin et al., 2014) – detected SMM with accuracies up to 99% (see Chapter 3 for additional details about previous classification work). However, there is very limited similar research for SIB classification. Previous work on SIB detection is limited to two reports that either created classifiers from trained actors imitating aggressive behaviors (Plötz et al., 2012), or focused on SMM with one example that could be considered SIB (Min, 2017). The former study extended models that were trained on imitated movement to one child with SIB, and found that classification with individual accelerometry data yielded accuracies on the order of 60-70% (Plötz et al., 2012). Classifiers may have had stronger performance if trained on natural data, versus simulated SIB, and should be tested on more than one participant with more than one behavior. One study also examined aggression, which may be related to SIB (Kanne & Mazurek, 2011). Naturally collected episodes of aggression were classified with high accuracy (AUC= 71-80 for individuals, AUC = 69 for group performance) using physiological and movement sensors, though SIB was not included in the activities of interest (Ozdenizci et al., 2018). Sensory aversions may preclude the use of physiological sensors for SIB (Y.-H. Chen et al., 2009), however, and so sensors not requiring skin contact were preferable for our application. Classification models in earlier studies were typically specific to each participant, with training and testing completed on each individual (Min, 2017; Plötz et al., 2012). When group-level models were employed, accuracy levels tended to decrease (e.g., from 80% for individual models to 69% for group-level models in (Ozdenizci et al., 2018). Further, classification methods used in earlier studies (e.g., SVM or neural networks) can have low interpretability, and other more accessible models should also be explored (e.g., regression, as in Ozdenizci et al., 2018). Multilevel regression models with varying intercepts and slopes may account for the variability among individual diagnoses of ASD and in presentations of behavior

(Gelman & Hill, 2006; Gowen & Hamilton, 2013), though such a model has yet to be applied to SIB.

Including additional features in group-level classifiers may provide superior performance to previously explored classification methods. Prior work classifying SMM in ASD used time- and frequency-domain features (Min et al., 2009), or focused on relatively simple measures of variability such as standard deviation and variance, which yielded high false positives (Coronato et al., 2014). In general, variability can be described using the linear measures included in SMM classification (Coronato et al., 2014), such as standard deviation, as well as using nonlinear dynamics measures, such as entropy (Robertson, Caldwell, Hamill, Kamen, & Whittlesey, 2013). Linear measures of variability are derived from descriptive methods and quantify the total variability within a system (Stergiou, 2004). For example, locomotion differences between severity levels of ASD have been quantified using the coefficient of variation, with increased stride variability found with increasing severity (Rinehart et al., 2006). However, when examining pathology, one method (e.g., linear) may show higher levels of variability compared to healthy controls and another method (e.g., nonlinear) may show lower levels (Bodfish, Parker, Lewis, Sprague, & Newell, 2001; Robertson et al., 2013). Variability of traditional linear measures may not align with nonlinear variability measures, as linear methods do not consider the time-related structure of data. Thus, both methods should be considered.

More effective features may be needed to capture the complexity of self-injurious movements, such as nonlinear measures derived from a dynamical systems perspective. Dynamical systems theory suggests that human movement changes and evolves over time, as governed by a deterministic process (Davids, Glazier, Araújo, & Bartlett, 2003). Nonlinear measures of variability, such as entropic and fractal measures, can quantify the evolution of movement over a period of time (Stergiou, 2004). Dynamical systems analyses separate variability components in a movement process, in particular to separate chaotic vs. deterministic variability (Stergiou, 2004). This theory may be relevant to SIB, since nonlinear components were found in temporal patterns of SIB, though varying within and between individuals (Sandman, Kemp, Mabini, Pincus, & Magnusson, 2012). More complex temporal patterns also emerged in the presence of SIB (Sandman et al., 2012), and recent work suggests that movements become increasingly complex as a child with ASD transitions to an episode of SIB (Kemp et al., 2016). This complexity may be captured through measuring nonlinear variability in movement of individuals with ASD and SIB. Nonlinear motor variability features could improve classification model performance by capturing this underlying variability in SIB movements, even if the SIB changes between episodes. Prior work has found that nonlinear movement measures, such as entropy, are indicative of diagnosis when applied to motor control in ASD. For example, children with ASD had decreased dynamical complexity during quiet stance compared to typically developing children (Bodfish et al., 2001; Fournier, Amano, Radonovich, Bleser, & Hass, 2014), although people with stereotypy showed greater linear variability (standard deviation) during postural sway (Bodfish et al., 2001). Similar to distinguishing between neurotypical and pathological movements, these methods may capture changes in pathology within an individual (i.e., detecting health changes such as early signs of aggression). Variability has also been associated with other pathological behaviors (Hausdorff et al., 2003; Robertson et al., 2013; Stergiou, 2004), and a progression of health conditions (Kalron, 2016; Socie, Motl, & Sosnoff, 2014), and thus could reflect changing risk of SIB in ASD as well. To our knowledge, though, only one study employed a nonlinear approach (recurrence quantification analysis, described in **Materials and methods**) to classifying motion in ASD, and found that the additional nonlinear features of motor variability improved classification accuracy by 5-9% (Großekathöfer et al., 2017). Though this analysis was performed on SMM, these results could generalize to SIB, which is similarly repetitive and rhythmic.

To our knowledge, smart monitoring for SIB has not yet been explored. The long-term goal of this research is to develop a real-time SIB monitoring system that can collect continuous movement data, alert the caregiver before SIB onset, and assist in management methods (e.g., redirecting the individual with SIB towards a different task). To this end, this study aimed to develop an interpretable model to identify a variety of behaviors among a range of participants, specifically by:

1) Applying dynamical systems theory in feature extraction to explore potential nonlinear motor variability in SIB movement, and
2) Implementing a logistic regression model with variable intercepts and slopes (multi-level logistic regression) to account for inter-individual variability.

Materials and methods

Participants

The data used here were obtained in the study described previously (Chapter 3). Children with SIB and ASD were recruited through the Virginia Tech Child Study Center and through the authors' networks. Caregivers were pre-screened to confirm several inclusion criteria: 1) children aged 5-14 years, reflecting heightened aggression in childhood (Hill et al., 2014; Kanne & Mazurek, 2011); 2) diagnosis of ASD; 3) SIB episodes >3/hour, to ensure multiple episodes during the 1-3 hour sessions; 4) fluency in English; and 5) home within driving distance of the Virginia Tech Child Study Center. Definitions of SIB vary (Lam & Aman, 2007; Rojahn, Matson, Lott, Esbensen, & Smalls, 2001). We selected a definition that reflects prior work (Iwata et al., 1994; Pace, Iwata, Edwards, & McCosh, 1986; Rojahn et al., 2001), and includes actions that could cause harm or pain to the self. SIB may serve different functions, excluding accidental incidents (e.g., tripping and falling) and cultural or aesthetic actions (e.g., ear piercings), and may or may not be performed with the intention to cause harm. SIB types may include classic behaviors (e.g., head banging, self-hitting or biting), fine behaviors (e.g., picking or pinching), and potentially injurious body contact with objects or the environment (e.g., repeatedly hitting into walls). In the current study, an SIB "episode" was defined as a continuous time period of a specific SIB type or of co-occurring SIB, potentially repetitive and rhythmic, preceded and followed by other types of SIB or other non-SIB actions.

Caregivers were asked during pre-screening to describe the details of their child's SIB, including known triggers, and preferred materials for sensors. We also requested that children wear clothing with pockets for sensor placement. We asked that caregivers bring the following items: toys for free play, items associated with triggering and managing SIB, and preferred clothes or accessories. All adult participants provided informed consent, and qualifying children (>7 years of age and of developmental level) provided assent before any data collection. The Virginia Tech Institutional Review Board approved all experimental procedures.

Eleven participants (5-14 years, M = 9.5, SD = 3.0) and their caregivers completed the study. Sessions lasted between 35-147 minutes, providing more than 1,000 minutes of data and >200 episodes of SIB. Ten of the 11 participants exhibited SIB (participants 1-4 and 6-11, denoted as "P#") with 18 different types (**Table 7**). Note that P1 wore two different sensor configurations (1a and 1b) in one session. P1 rejected pocket and ankle sensors during initial data collection, but accepted them before lower body SIB was shown in the latter half of the session. P8 rejected all sensors except wrist and pocket sensors, and P8 and P9 were siblings (each wore half of the sensors in consecutive sessions). All participants wore the wrist sensor, and the limited sensor configurations of P1, P8 and P9 precluded the use of other sensors in the group-level model. To include all participants in one group-level model, only the wrist sensor was considered. Data from 2-6 tri-axial accelerometers (**Table 7**) were input into the individual participant models.

Table 7. Participant identifier, type of SIB shown during the session, and sensors worn. The wrist sensor was commonly worn among all participants.

Participant #	**SIB**	**Sensors Configurations**
P1	Repeated foot to surface; Repeated fist to surface; Head hitting –with object	Wrist, waist (part 1) Wrist, waist, pockets, ankle (part 2)
P2	Finger picking; Scratching	Wrist, waist, pockets, ankles
P3	Heel to surface; Hand to surface	Wrist, waist, pockets, ankles
P4	Self-biting (hands, arms); Self-hitting; Pulling teeth; Eye-gauging; Jabbing pelvic region; Jabbing throat – location of prior tracheotomy; Hitting chin/jaw with heel of hand	Wrist, waist, pockets, ankles
P6	Repeatedly pulling on teeth using string/object; Foot to surface; Hand to surface	Wrist, waist, pockets, ankles
P7	Finger picking	Wrist, waist, pockets, ankles

P8	Foot to object; Hand to surface; Throwing body against object or surface	Wrists, pockets
P9	Finger picking; Lip picking; Head to wall	Wrist, waist, ankles
P10	Hands to surface; Finger picking; Scratching; Head to wall; Self-biting; Self-hitting; Eye-gauging; Pulling ear	Wrist, waist, pockets, ankles
P11	Finger picking; Hair pulling	Wrist, waist, pockets, ankles

Study overview

After obtaining consent at the noted Study Center, the lead author, or the caregiver guided by the author, secured sensors on the child where tolerated. Demographic information was obtained, including potential SIB triggers identified by the caregiver. A clinical psychology doctoral candidate then used well-established tools (ADOS and WASI, or Leite-R if the child was nonverbal) to confirm ASD diagnosis (Lord et al., 2012; Pugliese et al., 2015). Subsequently, movement sensors (see **Materials and methods**), video cameras, and 2-3 observing researchers monitored each child during free-play. Researchers instructed the caregivers to respond to SIB as if at home. If sufficient SIB episodes failed to occur during free play, caregivers had the option to prompt SIB in a controlled fashion with a commonly-used procedure (Standardized Observation Analogue Procedure, SOAP; Johnson et al., 2009). In this approach, caregivers chose everyday demands from a list of triggers (e.g., “Point to the toy”), or offered individual-specific demands that may elicit resistance and relate to SIB (e.g., “Wave to the researcher”). The noted doctoral candidate, overseen by a clinical psychologist (SWW), supervised SOAP to ensure prompt intervention when needed. Further details can be seen in Johnson et al. (2009) and in Chapter 3. The session ended when either: a) >3 episodes of SIB were observed, or b) participants or researchers stopped the session to prevent escalating behavior. At the end of the session, participants were compensated for their travel and time, and children selected a toy.

Instrumentation

Tri-axial accelerometers (ActiGraph GT9X Link, www.actigraphcorp.com) were used to track participant movement (sampling frequency = 60 Hz) throughout the session (**Figure 3**). Earlier work found that these particular sensors were both reliable and accurate when used with children and adolescents (Robusto & Trost, 2012) for tracking movement among both pathological and healthy populations (Bussmann et al., 2001). A maximum of six sensors were placed on the wrists, waist, pockets, and ankles as accepted by the participant. Sensor placement reflected prior research that found high accuracy when classifying activities with movement sensors on either the wrist or torso (Min et al., 2009; Trost, Zheng, & Wong, 2014). Ankle sensors were also included as potentially necessary to capture lower-body injurious behaviors (Plötz et al., 2012). Added materials (cotton, fleece, or sequins) were attached to the sensors as the participant desired (see Cantin-Garside et al., 2018 for a discussion of preferred materials). Sensor type, location, and attachment methods addressed caregiver needs determined in our previous study (Cantin-Garside et al., 2018), including flexible placement in discrete locations, removable, durable, and transferable across environments. Three Go-Pro™ cameras and an overhead camera recorded videos for "ground truth".

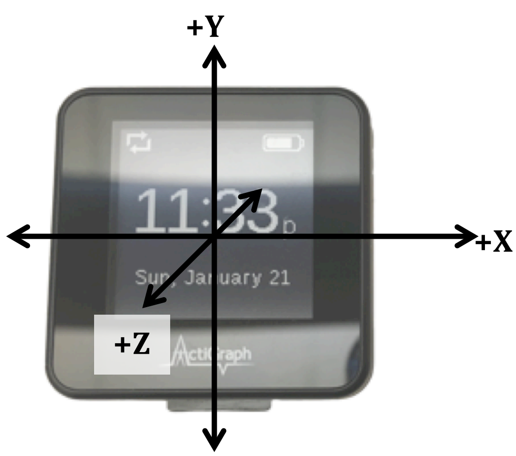


Figure 3. ActiGraph with labeled axes at initial placement.

Data processing and analysis

Sensor data were exported into MatLab R2018a (Classification Learner), which was used for data analysis and modeling (using an Intel dual-core 2.9 GHz CPU). Accelerometer data were labeled as non-SIB events (0) or SIB (1) using the ground truth video data and annotations from in-session observations. Before the session began, members of our research team discussed the SIB that parents described during pre-screening. Members of our team also discussed behaviors observed in-session both during and after the session. Behavioral definitions were further clarified prior to labeling data. Multiple researchers annotated and discussed the video data before labeling raw accelerometry files. (See **Figure 4** for a diagram of the modeling process).

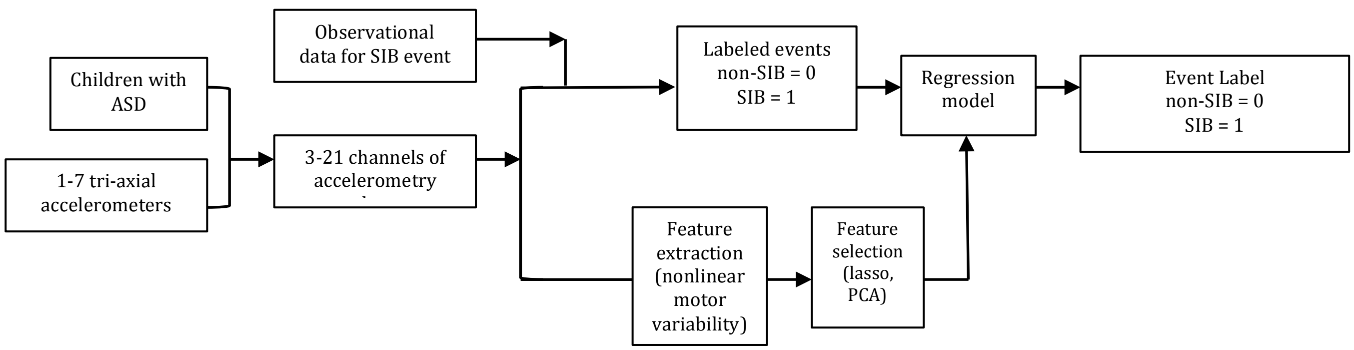


Figure 4. Overview of the data analysis and modeling process.

Raw sensor data were filtered using a 4th order, low-pass, recursive Butterworth filter with a cutoff frequency of 20Hz. Filtered data were used to obtain time- and frequency-domain features, whereas raw data were used for nonlinear motor variability feature extraction (see 2.5.1 for details) due to complications with nonlinearity and filtering described in prior work (Bruijn, Meijer, Beek, & Van Dieën, 2013; Mees & Judd, 1993). For continuous analysis of discrete data, all data were segmented into 2-second sliding windows with a 1 second overlap (Coronato et al., 2014; Goodwin et al., 2014). This short time window was use to minimize delays, which was considered important for real-time monitoring and reflects the potential for relatively short "bursts" of SIB.

Feature extraction

Three sets of features were extracted, including: 1) features in the time-domain; 2) features in the frequency-domain; and 3) nonlinear metrics from dynamical systems theory. As discussed above, the use of time- and frequency-domain features is supported by prior findings on classifying SMM (Coronato et al., 2014; Goncalves et al., 2012; Goodwin et al., 2014; Plötz et al., 2012), with nonlinear motor variability features included to capture dynamical complexity of motion and improve classifier performance (Samani, Srinivasan, Mathiassen, & Madeleine, 2015; Yentes et al., 2013). **Table 8** lists the features extracted for each channel. The presence (1) and absence (0) of a prompt to instigate SIB from SOAP procedures was also initially included during feature selection (see below).

Table 8. Features included three main types: time, frequency, and nonlinear motor variability. Features were calculated for each channel (XYZ), sensor, and participant, and reflect features in models classifying similar behaviors in prior work **(Coronato et al., 2014; Goncalves et al., 2012; Goodwin et al., 2014; Plötz et al., 2012)**

Feature Type	**Time Domain**	**Frequency Domain**	**Nonlinear Motor Variability**
Number of Features	19 features x 3 channels = 57 features	4 features x 3 channels = 12 features	9 features x 3 channels = 27 features
Features	• Channel cross-correlation coefficient • Mean difference between channels • Variance • Local Minima Count • Local Maxima Count • Peak • Minimum • Distribution percentages: 1, 10, 25, 50, 75, 90, 99 • Zero Crossings • Mean • Root mean square (RMS) • Jerk	• First two frequencies of FFT • First two corresponding amplitudes	• Detrended fluctuation analysis (DFA): · Exponent (α) • Entropy: · Sample entropy · Cross sample entropy • Recurrence Quantification Analysis Metrics (RQA): · Recurrence · Determinism · Laminarity · Divergence · Maximum diagonal length · Trapping time

Derivation of nonlinear metrics for motor variability

While deriving the time- and frequency-domain features was relatively straightforward, nonlinear motor variability metrics were extracted by first reconstructing the phase space of the raw sensor data (Stergiou, 2004). Phase space represents the states ("state space") of dynamical system behavior in a plot, and this reconstruction involves creating *M* copies of the original time series x, where *M* is the embedding dimension, using a time delay (τ) (Stergiou, 2004; Wurdeman, Myers, & Stergiou, 2013). The time delay was determined using two methods: 1) as the first minimum of the average mutual information function (Fraser & Swinney, 1986; Gates & Dingwell, 2009); 2) and the delay when the autocorrelation of the time series was less than e^{-1} (Samani et al., 2015; T. Rosenstein, J. Collins, & De Luca, 1993). Time delay was determined using both methods for SIB and separately for non-SIB events (see Samani et al., 2015, Stergiou, 2004, and Wurdeman et al., 2013 for additional details). The selected τ was the value for which both methods converged, and was similar for both SIB and non-SIB ($\tau = 5$). The resulting embedding dimension (*M*) was 4, derived from the global false nearest neighbor analysis method (Gates & Dingwell, 2009; Kennel, Brown, & Abarbanel, 1992). Both τ and *M* values here are similar to prior work on nonlinear variability of human motion (Gates & Dingwell, 2009). State space was reconstructed as embedding vectors $X(t)$ in the form of:

$$X(t) = [x(t), x(t+\tau), x(t+2\tau), \dots, x(t+(M-1)\tau)] \quad (1)$$

such that $i = 1, \dots, N-(M-1)$. Delay reconstruction was used to create the phase space, which produced consistent results in other work (Gates & Dingwell, 2009). The parameters described below were then calculated using the reconstructed phase space.

Entropy

Sample entropy (SaEn) was used to quantify the complexity of the acceleration signals, with low values indicating low complexity (Olofsen, W Sleigh, & Dahan, 2008; Richman & Moorman, 2000; Samani et al., 2015). SaEn was calculated using the following steps:

1. Compute $C_t^M(r) = \{number\ of\ X(i)\ such\ that\ ||X(t) - X(i)||_\infty \ll r\}$

where the tolerance threshold $r = 0.2 \cdot STDDEV(x)$ and $t \neq i$

2. Find $\phi^M(r) = \sum_{t=1}^{N-Md+1} \frac{C_t^M(r)}{N-Md+1}$

3. Calculate $SaEn = -ln \frac{\phi^{M+1}(r)}{\phi^M(r)}$

Entropy has been shown in prior studies to provide high accuracy when classifying SMM in ASD, though it has not yet been applied to SIB movement data (Albinali et al., 2011; Rad et al., 2016). Prior work has used SaEn as a reliable metric to capture changes over time in a short data set, since it is less sensitive than other nonlinear motor variability measures (e.g., approximate entropy) to changes in sample length (Yentes et al., 2013). Cross-sample entropy has not been explored in models classifying ASD motion, though was determined here between each sensor channel and used in our feature set. Cross sample entropy parallels SaEn in its calculation, but examines the difference between one data stream and another data stream (McCamley, Denton, Lyden, Yentes, & Medicine, 2017; Samani et al., 2015). Lower values imply similarity and synchronicity between the two data streams (Richman & Moorman, 2000).

Recurrence quantification analysis

Recurrence quantification analysis (RQA) was performed with a MatLab RQA toolbox (Chen & Yang, 2012; Yang, 2011) to evaluate phase space predictability and intermittency (Samani et al., 2015; Zbilut & Webber, 2006). An RQA map is first constructed through distance matrix comparison. A distance matrix (DM) consists of elements (DM_{ij}) that are Euclidian distances ($\mathrm{DM}_{ij} = d[X(i), X(j)]$) between embedding vectors $X(i)$ and $X(j)$ (Samani et al., 2015). DM_{ij} elements are then compared against a threshold determined by recurring dynamical trajectories, with elements = 1 for DM_{ij}< threshold, indicating recurrent points returned to a previous location, and = 0 otherwise (Großekathöfer et al., 2017; Samani et al., 2015). The selected threshold guarantees that the percentage of recurrent points remains within 0.1-2% of the total recurrent elements (Samani et al., 2015). RQA can be evaluated using several measures

(Richardson, Schmidt, & Kay, 2007; Samani et al., 2015; Zbilut & Webber, 2006), with the following selected as reliable for human subject research (McCamley et al., 2017; Richardson et al., 2007; Samani et al., 2015; Zbilut & Webber, 2006):

1) Recurrence – regularity of the time series as the percentage of recurrent points
2) Determinism – percentage of consecutive diagonally aligned recurrent points; relates to the inverse of the largest positive Lyapunov exponent because longer diagonals indicate deterministic versus chaotic movements
3) Laminarity – percentage of vertically aligned recurrent points (similar to determinism)
4) Divergence – inverse of the maximum diagonal line segment, related to the maximal Lyapunov exponent
5) Maximum diagonal length – proportional to the inverse of the maximal Lyapunov exponent
6) Trapping time – mean vertical line length indicating the duration of the trapped state

Detrended fluctuation analysis

Detrended fluctuation analysis (DFA) was used to quantify persistence of SIB movements. DFA exponents (α) were calculated for every time segment, to assess the long-range correlations (Dingwell & Cusumano, 2010; Stergiou, 2004), with persistence indicated by $0.5<\alpha <1$ for time series deviations that continue in the same direction, and anti-persistence by $0<\alpha< 0.5$ for deviations that continue in the opposite direction. DFA has been used in analyses of motor control for ASD, with evidence of long-range correlations (persistence) during a drawing task (Fleury, Kushki, Tanel, Anagnostou, & Chau, 2013). DFA has also been applied to capture the predictability of a movement, specifically walking (Dingwell & Cusumano, 2010). Persistence typically degrades with pathology; the underlying long-range correlations are altered in disordered movement, compared to consistent correlations in healthy movement, whether pathological behavior appears more restricted or more chaotic (Goldberger et al., 2002).

Feature selection

The least absolute shrinkage and selection operator (lasso) method was used to address multicollinearity, to remove redundant predictors, and determine the sparsest model when considering all features and all sensors (Tibshirani, 1996). This method directly selects variables that most contribute to the model. Principal components analysis (PCA) was then used for dimensional reduction through finding the optimal combination of the selected variables. A scree plot was examined to determine the optimal number of components when considering explained variance (Zwick & Velicer, 1986). Though the reliability of scree tests can be unstable, guidelines suggest it is more accurate and less variable than other component selection methods (Zwick & Velicer, 1986). Further, the number of included variables does not affect scree test accuracy, and results are interpretable compared to other component selection rules (Zwick & Velicer, 1986). With a large number of variables in our study and the exploratory nature of the included features, scree tests were used as a first effort toward selecting components to explain the variability in an SIB movement dataset. Variables with loadings >0.3 were considered as loaded on the PC (Cadima, 1995; Tabachnick & Fidell, 2013). This feature selection approach is capable of characterizing data despite high variability between participants. PCA output was subsequently used as input to a multilevel logistic regression model (MLR).

Regression modeling

A multilevel logistic regression (MLR) model was created with variable slopes and intercepts. The latter were used to account for high inter-subject variability (Gelman & Hill, 2006), which could be particularly relevant for ASD. Specifically, the model can be written as:

$$Y_i = logit^{-1}\left(\alpha_{j[i]} + \sum_k \beta_{j[i],k} X_k\right), i = 0,1, \dots I; j = 1,2, \dots, J; k = 1,2, \dots, K \qquad (2)$$

where i indexes over events, $j[i]$ is the index of a subject who exhibits event i, k indexes over the predictor variables, and $Y_i = 1$ is the outcome variable if the event is SIB and 0 otherwise. The intercept $\alpha_{j[i]}$ and the predictor slopes $\beta_{j[i],k}$ are variable for each subject. Both the intercepts $\alpha_{j[i]}$ and slopes of the predictors $\beta_{j[i],k}$ can be linearly modeled as:

$$\alpha_j \sim N\left(U_j \gamma_j^{\alpha}, \sigma_{\alpha}^2\right);\ \beta_{j,k} \sim N(V_{j,k}\gamma_{j,k}^{\beta}, \sigma_{\beta_k}^2) \qquad (3)$$

where U_j and $V_{j,k}$ are potential predictors, with corresponding linear coefficients γ_j' specific to the individual level, and modeling error variances σ^2.

Evaluation

Data were balanced and randomly selected following a 8:2 training/validation:testing ratio, so as to build a robust model and to test the built model (Poliker, 2006). SIB events are relatively rare compared to non-SIB events, which leads to a skewed distribution as found in prior work with ASD-related behaviors (Albinali, Goodwin, & Intille, 2009; Goodwin et al., 2014; Großekathöfer et al., 2017). SIB here lasted for about two seconds at minimum, though more subtle movements, such as picking, lasted longer, ranging from ~10-90 seconds. SIB and non-SIB data were balanced as in other work to address skewness (Albinali et al., 2009; Bastani, Kim, Kong, Nussbaum, & Huang, 2016; Bastani, Rao, & Kong, 2016; Großekathöfer et al., 2017; Rad et al., 2016). Balanced data were used for training, and 10-fold cross-validation (Albinali et al., 2011; Ozdenizci et al., 2018; Plötz et al., 2012). Two reserved and independent data sets were used for testing model generalization: 1) balanced data and 2) natural, unbalanced data reflecting the ratio of SIB:non-SIB in the complete dataset. All data were randomly selected from across the duration of a given session, and observations were assumed to reflect the entire dataset (Bastani, Kim, et al., 2016).

Outcome measures were calculated for each model (MLR, and the models described below) using true positives (A), false positives (B), false negatives (C), and true negatives (D) of each classification method (Webb, 2010). Classification performance was then quantified using the following metrics (Webb, 2010):

- Accuracy
  - (A+D)/(A+B+C+D)
- Specificity, or false positive rate
  - D/(B+D)
- Precision (“positive predictive value”)
  - A/(A+B)
- Recall (“sensitivity”)
  - A/(A+C)
- F-score
  - 2(Precision x Recall)/(Precision + Recall)

Training and testing time were also computed for all developed models to assess the potential for application to real-time monitoring.

Model comparisons

Additional group-level models were trained, validated, and tested, for the purpose of comparison with the MLR model ("MLR – variable intercepts and slopes"). These additional models were of five different types:

1. Logistic regression (LR) with variable intercept only ("LR - variable intercept"). This was used to compare MLR with a less complex model, while still accounting for participant variability.

2. LR without variable slopes or intercepts ("LR - no variable terms"). This model was included to compare MLR with a model that does not consider participant-level variation.

3. Two-way interaction model (stepwise LR), with included terms determined by BIC ("LR - stepwise"). This model was included to compare MLR with higher-order, nonlinear models, and to evaluate the effect on accuracy when including terms with lower interpretability but potentially higher predictive power.

4. Participant-level LR models ("LR-Ind"), one for each of the 10 who exhibited SIB. These models were included to compare group-level MLR with highly specific modeling that may have low generalizability, yet high accuracy.

5. Several models using machine learning algorithms shown to have high accuracy in related work: k-nearest neighbor – "kNN", with k=11 selected through optimization, support vector machines – "SVM", and decision trees – "DT"). These were included to compare MLR with previously employed methods demonstrating strong individual (though not group-level) performance in other ASD applications and our previous featureless work (Chapter 3), and to compare MLR with black-box models with lower interpretability but typically high predictive power.

Results

Dimensional reduction

Fifty-nine variables selected from lasso were included in PCA for the group-level MLR model, which included both linear and nonlinear motor variability features of each channel (**Appendix C, Table C1** for descriptive statistics of nonlinear motor variability). Lasso results did not include the prompt variable. PCA generated 12 principal components (PCs). Coefficients and explained variance of each PC are provided in **Appendix C, Table C2**. Over 23% of the variance was explained by PC1, with loadings primarily from frequency-based measures and measures capturing sudden or sharp movements (e.g., jerk, peak). Measures of the Z channel (vertical) loaded primarily on PC2, with coefficients >0.3 for mean absolute value and RMS. Nonlinear motor variability metrics had coefficients up to ~0.6 in selected PCs. Nonlinear metrics from RQA had coefficients >0.4 on PC6 (Z channel), >0.3 on PC8 (X channel), and >0.4 on PC9 (Y channel), while SaEn had coefficients >0.3 on PC10 and cross-sample entropy (YZ) >0.3 on PC12. The components listed above with nonlinear variable loadings >0.3 accounted for 11.1% of the total variance, (3.6%, 3.0%, and 2.6%, 2.5%, 1.9% respectively). All 12 components contributed to 65.6% of the group data variance.

Table 9 summarizes results using the MLR. PC1 and PC12 were both significant predictors when considered across participants, though not when varying with participants. The intercept, along with PC 2, 5, 6, 7 and 9, were only significant predictors when considering participant levels (Predictor|Par), and not when fixed. PCs 3, 8, 10, and 11 were significant predictors both when fixed and randomly varying with participant level. PC4 was not significant in the model either when fixed or when varying with participant level.

Table 9. Multi-level logistic regression parameter values for the group-level model including all 10 participants. Bold values indicate significant predictors in the model at the p<.05 level.

Parameter	Fixed Effect	Varying with Participant (Parameter|Par)
Intercept	-0.363	**1.192**
PC1	**0.033**	1.456e-15

PC2	-0.021	**0.217**
PC3	**0.247**	**0.274**
PC4	-0.001	1.712e-08
PC5	-0.031	**0.228**
PC6	0.060	**0.146**
PC7	0.003	**0.255**
PC8	**0.239**	**0.084**
PC9	0.028	**0.181**
PC10	**-0.273**	**0.185**
PC11	**0.333**	**0.250**
PC12	**0.132**	2.112e-09

Classifier performance

Tables 10 and **11** respectively summarize results regarding training time, accuracy, specificity, precision, recall, and F-scores of group-level models for validation and testing. Training times for MLR, stepwise LR, and cubic SVM were 2 - 4 magnitudes longer than for other classifiers (10^{-2} vs. 10^{-6} seconds/observation), though this same difference was not reflected in testing times (all times within 10^{-6} - 10^{-5} seconds/prediction). MLR had high accuracy (74.7%) and F-score (0.752) in validation, which decreased minimally when testing with balanced data (73.2% and 0.733). Accuracy and F-score decreased with unbalanced test data for MLR (69.1% and 0.184). Specificity, precision, and recall were all highest for MLR in validation (~0.73 - 0.77) and testing (~ 0.73 for all three measures). When considering participant levels with only a variable intercept versus both variable intercept and slopes, most performance metrics decreased by ~2-5%. LR without variable intercepts/slopes had the lowest accuracy (64.0%) in validation, dropping to 47.0% and 56.1% for balanced and unbalanced data, respectively. Linear SVM had the lowest specificity and precision (0.599 and 0.631), while LR without variable intercept/slopes had the lowest recall (0.663), though this trend did not extend to testing results. Stepwise LR had the lowest specificity for both balanced and unbalanced test data (0.526 and 0.552, respectively). LR without variable intercepts/slopes had the lowest precision, recall, and F-score for balanced test data (0.455, 0.304, and 0.365 respectively). The kNN classifier had the lowest precision, recall, and F-score for unbalanced test data (0.064, 0.591 and 0.116, respectively), while LR without variable intercept/slopes had the lowest F-score for validation (0.648) and balanced test data (0.365).

Table 10. Validation results for group-level classifiers.

Classifier	Training Sec/Observation	Accuracy	Specificity	Precision	Recall	F
MLR - variable intercept and slopes	1.49E-02	0.747	0.728	0.738	0.766	0.752
LR - variable intercept	4.71E-04	0.705	0.676	0.694	0.734	0.713
LR - No variable terms	8.34E-06	0.640	0.617	0.634	0.663	0.648
LR - stepwise	8.49E-02	0.671	0.673	0.672	0.669	0.670
kNN, k = 11	1.49E-05	0.676	0.621	0.659	0.731	0.693
SVM - Linear	1.66E-04	0.642	0.599	0.631	0.685	0.657
SVM - Cubic	6.78E-03	0.696	0.657	0.682	0.734	0.707
SVM - Gaussian	1.28E-04	0.690	0.661	0.679	0.719	0.699
DT	9.94E-06	0.683	0.652	0.672	0.713	0.692

Table 11. Test results at the group level.

Algorithm	Test Type	Prediction Sec/ Observation	Accuracy	Specificity	Precision	Recall	F
MLR - variable intercept and slopes	1	5.98E-05	0.732	0.729	0.731	0.735	0.733
	2	5.70E-05	0.691	0.687	0.105	0.773	0.184

LR - variable intercept	1	1.54E-05	0.705	0.696	0.702	0.715	0.708
	2	1.66E-05	0.647	0.641	0.092	0.773	0.165
LR - No variable intercept/slopes	1	1.33E-05	0.470	0.637	0.455	0.304	0.365
	2	1.07E-05	0.561	0.552	0.073	0.750	0.134
LR - Stepwise	1	1.45E-05	0.488	0.526	0.487	0.450	0.467
	2	1.45E-05	0.561	0.552	0.073	0.750	0.134
kNN, k = 11	1	2.79E-05	0.643	0.567	0.624	0.719	0.668
	2	2.00E-05	0.591	0.591	0.064	0.591	0.116
SVM - Linear	1	5.11E-05	0.619	0.612	0.617	0.626	0.622
	2	4.60E-05	0.551	0.542	0.072	0.750	0.131
SVM - Cubic	1	4.76E-05	0.677	0.639	0.664	0.715	0.688
	2	4.67E-05	0.625	0.622	0.081	0.705	0.145
SVM - Gaussian	1	4.68E-05	0.671	0.641	0.662	0.702	0.681
	2	4.09E-05	0.645	0.646	0.076	0.614	0.135
Decision Tree	1	1.14E-05	0.695	0.676	0.688	0.715	0.701
	2	9.86E-06	0.640	0.638	0.082	0.682	0.146

Tables 12 and **13** show the classifier performance for validating and testing individual models, respectively (see **Appendix C** for sample graphs of features and labels). Computational time was on the order of 10^{-5}-10^{-3} seconds/observation for training and 10^{-4}-10^{-3} seconds/prediction for testing. Validation accuracy was higher overall (70.7-97.9%) when compared to group-level models (64.0-74.7%). Testing accuracy ranged widely, from 50.0-83.3% for balanced data to 27.2-95.7% for unbalanced data. Likewise, specificity ranged from 0.648-1 for validation datasets, and from 0.500-1 for balanced and unbalanced test data. Precision was between 0.692-1 for validation data, 0-0.818 for balanced test data, and 0-0.5 for unbalanced test data. The ranges of recall values were 0.746-1 for validation data and 0-1 for both balanced and unbalanced test data. F-scores ranged from 0.718-0.979 for validation data, 0-0.857 for balanced test data, and 0-0.667 for unbalanced data.

Table 12. Validation results for individual participants. Note that 1a = first part of P1 session with only upper body sensors, and 1b = second part of P1 session with additional lower body sensors.

P	Training Sec/Observation	Validation Accuracy	Specificity	Precision	Recall	F-score
1a	2.30E-03	0.893	0.857	0.867	0.929	0.897
1b	1.86E-03	0.979	1.000	1.000	0.958	0.979
2	1.61E-03	0.883	0.878	0.879	0.888	0.883
3	1.11E-03	0.803	0.770	0.785	0.836	0.810
4	7.84E-05	0.707	0.668	0.692	0.746	0.718
6	2.61E-04	0.857	0.838	0.844	0.876	0.860
7	1.55E-04	0.760	0.648	0.713	0.872	0.784
8	2.19E-03	0.935	0.870	0.885	1.000	0.939
9	1.74E-04	0.941	0.929	0.931	0.953	0.941
10	1.25E-04	0.772	0.797	0.786	0.747	0.766
11	2.24E-04	0.933	0.922	0.924	0.943	0.933

Table 13. Test results for individual participants. Note that 1a = first part of P1 session with only upper body sensors, and 1b = second part of P1 session with additional lower body sensors.

P	Test Type	Prediction Sec/Prediction	Accuracy	Specificity	Precision	Recall	F-score
1a	1	4.56E-03	0.833	0.667	0.750	1	0.857
1a	2	3.89E-03	0.833	0.800	0.500	1	0.667
1b	1	1.62E-03	0.500	1	0	0	0
1b	2	1.75E-03	0.917	1	0	0	0
2	1	2.94E-04	0.500	1	0	0	0
2	2	6.23E-04	0.938	1	0	0	0
3	1	1.80E-03	0.733	0.800	0.769	0.667	0.714
3	2	5.28E-04	0.533	0.517	0.067	1	0.125
4	1	1.04E-04	0.687	0.613	0.663	0.761	0.708
4	2	1.08E-04	0.577	0.525	0.263	0.833	0.400
6	1	1.66E-04	0.828	0.810	0.817	0.845	0.831
6	2	1.72E-04	0.543	0.514	0.117	1	0.209
7	1	2.97E-04	0.779	0.706	0.744	0.853	0.795
7	2	2.05E-04	0.272	0.238	0.057	1	0.108

8	1	2.30E-03	0.500	0.500	0.500	0.500	0.500
8	2	2.43E-03	0.538	0.500	0.143	1	0.250
9	1	2.76E-04	0.500	1	0	0	0
9	2	1.86E-04	0.925	1	0	0	0
10	1	1.76E-04	0.767	0.787	0.778	0.747	0.762
10	2	1.56E-04	0.673	0.655	0.140	1	0.246
11	1	3.67E-04	0.800	0.829	0.818	0.771	0.794
11	2	4.60E-04	0.957	1	0	0	0

Discussion

In this study, an interpretable model was developed to identify diverse types of SIB among a range of participants. Nonlinear motor variability features from dynamical systems theory contributed to a multi-level logistic regression model capable of detecting SIB at the group level, with selected components from dimension reduction explaining >65% of the variance in the dataset. The Lasso method did not select the prompt variable for this group-level model (recall that this prompt represented the presence or absence of caregiver actions that were targeted at instigating SIB), indicating that this model explains the presence of SIB beyond an identified SIB trigger. This finding is consistent with a prior report showing that temporal patterns in SIB occur independent of behavioral or environmental influences (or “triggers”; Sandman et al., 2012). Instead, nonlinear motor variability features (e.g., from RQA and entropy) loaded on PCs that accounted for >10% of the explained variance in the dataset. DFA features had more moderate loadings on PCs, and these features were a novel addition to modeling for ASD. Descriptive statistics of nonlinear motor variability metrics from pooled SIB versus non-SIB events across participants show little difference between the behavioral classes (Appendix C, Table C1), which may explain the poor performance of a general group-level model without participant levels. However, upon examining one of the most severe behaviors (head banging) in one participant, there are clear differences in the nonlinear motor variability of SIB versus non-SIB events. For example, DFA exponents for both non-SIB and SIB events in the Y and Z axes are slightly anti-persistent (<0.5), indicating changes evolving in different directions over time. Though exponents remain < 0.5, there is a slight increase in DFA exponents for the Y and Z axis for SIB events compared to non-SIB, indicating less anti-persistence in SIB than non-SIB events.

Differences among other nonlinear metrics were evident for head banging in P9. Recurrence rate, for example, decreased for head banging in P9 when compared to non-SIB events; this finding indicates a lower level of regularity in SIB data. This finding opposes the commonly accepted perception that repetitive behaviors are "regular". There was a slight decrease in sample entropy for SIB in the Y and Z axis when compared to non-SIB events, suggesting lower levels of complexity; however, cross-sample entropy increased during SIB, indicating higher levels of complexity between two channels of data. These findings may indicate that SIB occurs due to over/understimulation to seek system stability ("less" or "more" complexity). Together, these findings suggest underlying nonlinear trends in the movements occurring during SIB. Additional variance in PCA was explained by classic time and frequency domain features, similar to prior work that used time and frequency features to accurately classify SMM (Coronato et al., 2014; Goncalves et al., 2012; Goodwin et al., 2014; Plötz et al., 2012). LR without variable intercept or slopes performed inferiorly to LR with a variable intercept, and MLR with both variable intercept and slopes. MLR with both variable intercept and slopes performed superior to all other classifiers, including commonly used machine learning algorithms implemented in other work (Albinali et al., 2011; Goodwin et al., 2014). These findings imply that inter-individual variability also contributed to dataset variance.

This study is the first, to our knowledge, to use nonlinear metrics to explain the variance in SIB movement data. Our findings suggest that movements in SIB can be described as a dynamical system with long-term deviations, which is consistent with prior evidence that stereotypical motor movements in ASD can be accurately detected using nonlinear features from RQA (Großekathöfer et al., 2017). Similarly, our feature extraction revealed higher loadings for RQA features when compared to other nonlinear factors, and the associated principal components were significant in our MLR model. Our findings, along with those of (Großekathöfer et al., 2017), suggest that RQA metrics may be critical in detecting repetitive and rhythmic motor movements (such as stereotypical motor movements, and SIB) in ASD. Further, all PC with nonlinear variable loadings >0.1 were significant in the MLR model. PC6 and PC9, with loadings primarily from RQA metrics from the Z and Y axis, respectively, were significant in the model only when randomly varying with participant level; these metrics are not sufficient to classify identify SIB unless considering variable slopes, which indicates that nonlinear movement aspects may be specific to each individual. Time- and frequency-domain features that loaded on PC1 (frequency components, jerk, and peak/minimum) were significant predictors in MLR when independent from participant levels, indicating that these variables may be predictive of SIB without considering individual variability; the only nonlinear metric of motor variability that this finding applied to was PC12 (cross-sample entropy). Other PCs with loadings from nonlinear motor variability metrics (RQA for X on PC8, sample entropy on PC10) were significant when considered as either a fixed or a variable effect, implying that features such as sample entropy vary consistently between participants while still explaining individual-level variability. Nonlinear measures were significant predictors of SIB when they varied with participants, which supports prior work that found nonlinear components of movement were specific to individuals (Elizabeth B Torres & Donnellan, 2015).

We believe the current group-level model is the first to achieve accuracy ~75% when identifying SIB among a diverse group of behaviors and participants. Previous research on classifying other repetitive motor movements (Albinali et al., 2011; Coronato et al., 2014; Goncalves et al., 2012; Goodwin et al., 2014; Goodwin et al., 2008; Großekathöfer et al., 2017; Min, 2017; Min et al., 2009; Rad et al., 2016) and aggression (Ozdenizci et al., 2018; Plötz et al., 2012) has evaluated specific models trained and/or tested on individual participants only, and performance dropped from 80% with individual models to 69% when applied to the group of participants (Ozdenizci et al., 2018). Though a similar decrease was also evident here, percent accuracy was ~6% higher than earlier group-level accuracy percentages. This increased classifier performance may be due, at least in part, from the use of feature selection and dimensional reduction methods, along with the multi-level properties of our model that accounted for inter-individual variability. Also, a larger sample of participants was included here, compared to earlier modeling reports of ASD behaviors, with samples ranging from one to six (Albinali et al., 2011; Coronato et al., 2014; Goncalves et al., 2012; Goodwin et al., 2014; Goodwin et al., 2011; Min, 2017; Min et al., 2009). Further, our participant pool encompassed 18 different behaviors across children 5-14 years of age, suggesting potential generality to a wider sample of children with ASD and SIB.

As in the work of (Ozdenizci et al., 2018), regression showed promising results here compared to other classifiers; however, the earlier authors focused on aggression towards others, whereas our study applied regression on SIB data. MLR here had higher accuracy, specificity, precision, and recall compared to several commonly-used machine learning algorithms. The selected machine learning algorithms also detected SIB with high accuracy in our previous study using featureless data, though accuracy greatly decreased at the group-level (Chapter 3). Multi-level regression with both variable slopes and intercept may be preferred for group data with variable behaviors that could be specific to an individual, and it may also be more accessible to interpretation than other machine learning algorithms. MLR had classifier performance superior to LR without varying intercept/slope, further emphasizing the potential importance of individualizing models for participants with ASD and SIB. MLR, though, was only inferior to several highly-specific participant models, which may not generalize beyond the participant. The fact that we found widely varying performance metrics across participants, however, may account for the unexplained variance in MLR. Several participants had either near perfect detection (e.g., P1) or quite poor detection (P8 or P9) in testing. This wide range could have resulted from the inconsistent amount of SIB data included (P9 had the shortest session of all participants), the different types of SIB, or the variable sensor configurations between participants. Of note, MLR only included the one sensor commonly worn by all participants: the wrist sensor. This single wrist sensor may not be sufficient for all SIB types, such as head banging or kicking, and therefore may have led to decreased performance measures in the group model. Yet, despite only having one sensor to incorporate in the group model in this study, MLR still showed superior performance to all other tested classifiers. Group-level classifiers may be more practical (i.e., efficient and generalizable) to implement in real-world applications, and the current results are promising for automatically identifying diverse SIB with minimally-invasive technology.

Limitations and future work

This study provides initial groundwork toward creating a group-level classifier for SIB classification, yet further research is needed in several respects. Such work should include expanded data to improve classifier performance and extend the model for more individuals with ASD, given the highly heterogeneous ASD diagnoses and the lifelong pervasiveness of ASD and SIB. SIB presentation can be extremely variable, including in its duration. SIB data here may not have been sufficient for some participants, such as P8, when SIB episodes are few and/or short (<100 seconds). There would thus likely be value in monitoring SIB across several days, to capture additional episodes across different contexts. Data from other episodes may also help increase explanatory power for the MLR, though accuracy is perhaps more critical for online detection of SIB. It may be possible to improve the MLR by using nonlinear terms with variable slopes, though this could decrease interpretability of the model. Additional levels could also be incorporated into the model, such as age or SIB type.

Our work supports the presence of nonlinear motor variability within SIB. However, several nonlinear motor variability features (DFA) loaded only modestly ($< \sim 0.1$) on PCs, which may be due to the young ages of the participant pool. The long-range correlations quantified by DFA only develop in gait during childhood (Ary L Goldberger et al., 2002), so such correlations may not yet be evident in SIB movements if the participants are young. Older participants may show more explicit anti/persistence in pathological movement, in contrast, which could lead to additional evidence of nonlinear motor variability in SIB. Dynamic movement signatures of individuals with ASD could provide information to detect pathology, such as movements involved in SIB, before typical diagnostic measures (Elizabeth B Torres & Donnellan, 2015), and could explain the pattern of SIB onset. These individual movement signatures might also reveal trends about intentions that underlie SIB movements, such as whether the motion is goal-directed or spontaneous (Elizabeth B Torres & Donnellan, 2015), or the etiology of ASD, through mapping movement characteristics to underlying mechanisms of movement (Rinehart et al., 2006). With additional information about ASD movement signatures, variability components (quantified with metrics such as RQA and entropy) could be the basis for an intervention to promote self-awareness and intentional movements in ASD (Torres, Yanovich, & Metaxas, 2013).

In future work, we plan to use our current findings to build more sophisticated hierarchical models. One useful addition might be including Bayesian priors for individual-specific information. If such models improve accuracy while retaining interpretability, they could be used to determine the necessity of intervention at the earliest indicator of the event. Specifically, the predicted probability score from logistic regression can serve as a criterion for caregiver interjection, by setting a pre-determined threshold (i.e., caregivers should interject if the probability of SIB > 0.8). A monitoring system could also include real-time estimation of variable parameters (intercept and slope) for each individual with ASD and SIB. Parameter coefficients can be estimated by the empirical Bayes approach, which allows the mean value of the prior distribution to equal the mean of coefficients from the training data. Using new data in real-time, the posterior distribution can then be recomputed and updated for that participant. Continuously adapting predictors could both improve current models and address evolving behavior when tracking SIB.

Conclusions

This work provides a framework for – and initial results obtained from – interpretable SIB classification at the group-level, particularly through introducing new predictors with variable slopes and intercepts in a multi-level classifier. A new application of nonlinear metrics to movement in SIB was employed, specifically to develop a group-level classification model. We found that both linear/nonlinear measures of motor variability and time/frequency-domain features, paired with feature selection and dimensional reduction, explained >65% of the variance found in SIB movement data, and classified diverse SIB types among a group of 10 participants with ~75% accuracy. Our results are promising in terms of the feasibility of developing a continuous monitoring system for SIB that can be applied to different types of behaviors and a range of individuals. Future work should continue to build on these results, with added consideration of prior distributions for adaptive modeling.

Acknowledgements

This research was supported by a National Science Foundation Graduate Research Fellowship (to the author) and an Institute of Society, Culture and Environment (ISCE) Research Grant (to DS). However, neither agency had any involvement in data analysis, interpretation, or the decision to publish.

References

Albinali, F., Goodwin, M. S., & Intille, S. (2011). Detecting stereotypical motor movements in the classroom using accelerometry and pattern recognition algorithms. *Pervasive and Mobile Computing, 8*(1), 103. doi:10.1016/j.pmcj.2011.04.006

Albinali, F., Goodwin, M. S., & Intille, S. S. (2009). *Recognizing stereotypical motor movements in the laboratory and classroom: A case study with children on the autism spectrum.*

Paper presented at the Proceedings of the 11th international conference on Ubiquitous computing.
Allen, K. D., & Warzak, W. J. (2000). The problem of parental nonadherence in clinical behavior analysis: Effective treatment is not enough. *Journal of Applied Behavior Analysis, 33*(3), 373-391. doi:10.1901/jaba.2000.33-373
APA. (2013). *Dsm 5*: American Psychiatric Association.
Baio, J., Wiggins, L., Christensen, D. L., Maenner, M. J., Daniels, J., Warren, Z., . . . White, T. (2018). Prevalence of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 sites, united states, 2014. *MMWR Surveillance Summaries, 67*(6), 1.
Bastani, K., Kim, S., Kong, Z. J., Nussbaum, M. A., & Huang, W. (2016). Online classification and sensor selection optimization with applications to human material handling tasks using wearable sensing technologies. *IEEE Transactions on Human-Machine Systems, 46*(4), 485-497. doi:10.1109/THMS.2016.2537747
Bastani, K., Rao, P. K., & Kong, Z. (2016). An online sparse estimation-based classification approach for real-time monitoring in advanced manufacturing processes from heterogeneous sensor data. *IIE Transactions, 48*(7), 579. doi:10.1080/0740817X.2015.1122254
Bellini, S., & Akullian, J. (2007). A meta-analysis of video modeling and video self-modeling interventions for children and adolescents with autism spectrum disorders. *Exceptional Children, 73*(3), 264-287.
Bodfish, J. W., Parker, D. E., Lewis, M. H., Sprague, R. L., & Newell, K. M. (2001). Stereotypy and motor control: Differences in the postural stability dynamics of persons with stereotyped and dyskinetic movement disorders. *American Journal on Mental Retardation, 106*(2), 123-134.
Bruijn, S., Meijer, O., Beek, P., & Van Dieën, J. (2013). Assessing the stability of human locomotion: A review of current measures. *Journal of the Royal Society Interface, 10*(83), 20120999.
Bussmann, J. B. J., Martens, W. L. J., Tulen, J. H. M., Schasfoort, F. C., van den Berg-Emons, H. J. G., & Stam, H. J. (2001). Measuring daily behavior using ambulatory accelerometry: The activity monitor. *Behavior Research Methods, Instruments, & Computers, 33*(3), 349-356. doi:10.3758/bf03195388
Cabibihan, J.-J., Javed, H., Aldosari, M., Frazier, T., & Elbashir, H. (2017). Sensing technologies for autism spectrum disorder screening and intervention. *Sensors, 17*(1), 46.
Cadima, J., & Jolliffe, I. T. J. J. o. A. S. (1995). Loading and correlations in the interpretation of principle compenents. *22*(2), 203-214.
Cantin-Garside, K., Valdez, R. S., Nussbaum, M. A., White, S., Kim, S., Kim, C. D., & Fortes, D. M. (2018). *Exploring challenges of monitoring technology and self-injurious behavior in autism spectrum disorder.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.
Chen, Y., & Yang, H. (2012). Multiscale recurrence analysis of long-term nonlinear and nonstationary time series. *Chaos, Solitons & Fractals, 45*(7), 978-987. doi:https://doi.org/10.1016/j.chaos.2012.03.013
Chen, Y.-H., Rodgers, J., & McConachie, H. (2009). Restricted and repetitive behaviours, sensory processing and cognitive style in children with autism spectrum disorders.

Journal of Autism and Developmental Disorders, 39(4), 635-642. doi:10.1007/s10803-008-0663-6
Coronato, A., De Pietro, G., & Paragliola, G. (2014). A situation-aware system for the detection of motion disorders of patients with autism spectrum disorders. *Expert Systems with Applications, 41*(17), 7868-7877.
Davids, K., Glazier, P., Araújo, D., & Bartlett, R. (2003). Movement systems as dynamical systems. *Sports medicine, 33*(4), 245-260.
Dawson, G., Rogers, S., Munson, J., Smith, M., Winter, J., Greenson, J., . . . Varley, J. (2010). Randomized, controlled trial of an intervention for toddlers with autism: The early start denver model. *Pediatrics, 125*(1), e17-e23.
Dingwell, J. B., & Cusumano, J. P. (2010). Re-interpreting detrended fluctuation analyses of stride-to-stride variability in human walking. *Gait & Posture, 32*(3), 348-353. doi:10.1016/j.gaitpost.2010.06.004
Dracobly, J. D., Dozier, C. L., Briggs, A. M., & Juanico, J. F. (2018). Reliability and validity of indirect assessment outcomes: Experts versus caregivers. *Learning and Motivation, 62*, 77-90. doi:https://doi.org/10.1016/j.lmot.2017.02.007
Dunlap, G., Newton, J. S., Fox, L., Benito, N., & Vaughn, B. (2001). Family involvement in functional assessment and positive behavior support. *Focus on autism and other developmental disabilities, 16*(4), 215-221.
Fleury, A., Kushki, A., Tanel, N., Anagnostou, E., & Chau, T. (2013). Statistical persistence and timing characteristics of repetitive circle drawing in children with asd. *Developmental Neurorehabilitation, 16*(4), 245-254. doi:10.3109/17518423.2012.758184
Fournier, K. A., Amano, S., Radonovich, K. J., Bleser, T. M., & Hass, C. J. (2014). Decreased dynamical complexity during quiet stance in children with autism spectrum disorders. *Gait & Posture, 39*(1), 420-423. doi:10.1016/j.gaitpost.2013.08.016
Fraser, A. M., & Swinney, H. L. (1986). Independent coordinates for strange attractors from mutual information. *Physical review A, 33*(2), 1134.
Gates, D. H., & Dingwell, J. B. (2009). Comparison of different state space definitions for local dynamic stability analyses. *J Biomech, 42*(9), 1345-1349. doi:http://dx.doi.org/10.1016/j.jbiomech.2009.03.015
Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*: Cambridge university press.
Goldberger, A. L., Amaral, L. A., Hausdorff, J. M., Ivanov, P. C., Peng, C.-K., & Stanley, H. E. (2002). Fractal dynamics in physiology: Alterations with disease and aging. *Proceedings of the National Academy of Sciences, 99*(suppl 1), 2466-2472.
Goldberger, A. L., Amaral, L. A. N., Hausdorff, J. M., Ivanov, P. C., Peng, C.-K., & Stanley, H. E. (2002). Fractal dynamics in physiology: Alterations with disease and aging. *99*(suppl 1), 2466-2472. doi:10.1073/pnas.012579499 %J Proceedings of the National Academy of Sciences
Goncalves, N., Rodrigues, J. L., Costa, S., & Soares, F. (2012). *Preliminary study on determining stereotypical motor movements*, United States.
Goncalves, N., Rodrigues, J. L., Costa, S., & Soares, F. (2012, 2012). *Automatic detection of stereotyped hand flapping movements: Two different approaches*.
Goodwin, M., Haghighi, M., Tang, Q., Akcakaya, M., Erdogmus, D., & Intille, S. (2014, 2014). *Moving towards a real-time system for automatically recognizing stereotypical motor movements in individuals on the autism spectrum using wireless accelerometry*.

Goodwin, M. S., Intille, S. S., Albinali, F., & Velicer, W. F. (2011). Automated detection of stereotypical motor movements. *Journal of Autism and Developmental Disorders, 41*(6), 770-782. doi:10.1007/s10803-010-1102-z

Goodwin, M. S., Intille, S. S., Velicer, W. F., & Groden, J. (2008). *Sensor-enabled detection of stereotypical motor movements in persons with autism spectrum disorder.* Paper presented at the Proceedings of the 7th international conference on Interaction design and children.

Gowen, E., & Hamilton, A. (2013). Motor abilities in autism: A review using a computational context. *Journal of Autism and Developmental Disorders, 43*(2), 323. doi:10.1007/sl0803-012-1574-0

Großekathöfer, U., Manyakov, N. V., Mihajlović, V., Pandina, G., Skalkin, A., Ness, S., . . . Goodwin, M. S. (2017). Automated detection of stereotypical motor movements in autism spectrum disorder using recurrence quantification analysis. *Frontiers in Neuroinformatics, 11*, 9. doi:10.3389/fninf.2017.00009

Hausdorff, J., Schaafsma, J., Balash, Y., Bartels, A., Gurevich, T., & Giladi, N. (2003). Impaired regulation of stride variability in parkinson's disease subjects with freezing of gait. *Experimental Brain Research, 149*(2), 187-194.

Hill, A. P., Zuckerman, K. E., Hagen, A. D., Kriz, D. J., Duvall, S. W., van Santen, J., . . . Fombonne, E. (2014). Aggressive behavior problems in children with autism spectrum disorders: Prevalence and correlates in a large clinical sample. *Research in Autism Spectrum Disorders, 8*(9), 1121-1133. doi:10.1016/j.rasd.2014.05.006

Iwata, B. A., Pace, G. M., Dorsey, M. F., Zarcone, J. R., Vollmer, T. R., Smith, R. G., . . . Willis, K. D. (1994). The functions of self-injurious behavior: An experimental-epidemiological analysis. *27*(2), 215-240. doi:doi:10.1901/jaba.1994.27-215

Johnson, C. R., Butter, E. M., Handen, B. L., Sukhodolsky, D. G., Mulick, J., Lecavalier, L., . . . McDougle, C. J. (2009). Standardised observation analogue procedure (soap) for assessing parent and child behaviours in clinical trials. *J Intellect Dev Disabil, 34*(3), 230-238. doi:10.1080/13668250903074471

Kalron, A. (2016). Gait variability across the disability spectrum in people with multiple sclerosis. *Journal of the Neurological Sciences, 361*, 1-6. doi:http://dx.doi.org/10.1016/j.jns.2015.12.012

Kanne, S. M., & Mazurek, M. O. (2011). Aggression in children and adolescents with asd: Prevalence and risk factors. *Journal of Autism and Developmental Disorders, 41*(7), 926-937. doi:10.1007/s10803-010-1118-4

Kemp, A. S., Lenjavi, M. R., Touchette, P. E., Pincus, D., Magnusson, M. S., & Sandman, C. A. (2016). The self-organization of self-injurious behavior as revealed through temporal pattern analyses. *Discovering Hidden Temporal Patterns in Behavior and Interaction: T-Pattern Detection and Analysis with THEME™*, 101-124.

Kennel, M. B., Brown, R., & Abarbanel, H. D. (1992). Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Physical review A, 45*(6), 3403.

Kirby, A. V., Boyd, B. A., Williams, K. L., Faldowski, R. A., & Baranek, G. T. (2016). Sensory and repetitive behaviors among children with autism spectrum disorder at home. *Autism*, 1-13.

Krahn, G. L., Klein Walker, D., & Correa-De-Araujo, R. (2015). Persons with disabilities as an unrecognized health disparity population. *American Journal of Public Health, 105*(S2), S198-S206. doi:10.2105/AJPH.2014.302182

Kurtz, P. F., Chin, M. D., Huete, J. M., Tarbox, R. S., O'Connor, J. T., Paclawskyj, T. R., & Rush, K. S. (2003). Functional analysis and treatment of self ‐ injurious behavior in young children: A summary of 30 cases. *Journal of Applied Behavior Analysis, 36*(2), 205-219.

Lam, K. S., & Aman, M. G. (2007). The repetitive behavior scale-revised: Independent validation in individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 37*(5), 855-866.

Liptak, G. S., Benzoni, L. B., Mruzek, D. W., Nolan, K. W., Thingvoll, M. A., Wade, C. M., & Fryer, G. E. (2008). Disparities in diagnosis and access to health services for children with autism: Data from the national survey of children's health. *J Dev Behav Pediatr, 29*(3), 152-160. doi:10.1097/DBP.0b013e318165c7a0

Lord, C., Rutter, M., DiLavore, P., Risi, S., Gotham, K., & Bishop, S. (2012). Autism diagnostic observation schedule–2nd edition (ados-2). *Los Angeles, CA: Western Psychological Corporation*.

Lucyshyn, J. M., Albin, R. W., Horner, R. H., Mann, J. C., Mann, J. A., & Wadsworth, G. (2007). Family implementation of positive behavior support for a child with autism longitudinal, single-case, experimental, and descriptive replication and extension. *Journal of Positive Behavior Interventions, 9*(3), 131-150.

Marcu, G., Tassini, K., Carlson, Q., Goodwyn, J., Rivkin, G., Schaefer, K. J., . . . Kiesler, S. (2013). *Why do they still use paper?: Understanding data collection and use in autism education*. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Paris, France.

McCamley, J., Denton, W., Lyden, E., Yentes, J. M. J. C., & Medicine, M. M. i. (2017). Measuring coupling of rhythmical time series using cross sample entropy and cross recurrence quantification analysis. *2017*.

McTiernan, A., Leader, G., Healy, O., & Mannion, A. (2011). Analysis of risk factors and early predictors of challenging behavior for children with autism spectrum disorder. *Research in Autism Spectrum Disorders, 5*(3), 1215-1222. doi:https://doi.org/10.1016/j.rasd.2011.01.009

Mees, A. I., & Judd, K. (1993). Dangers of geometric filtering. *Physica D: Nonlinear Phenomena, 68*(3-4), 427-436.

Min, C.-H. (2017). *Automatic detection and labeling of self-stimulatory behavioral patterns in children with autism spectrum disorder.* Paper presented at the Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE.

Min, C.-H., Tewfik, A. H., Kim, Y., & Menard, R. (2009, 2009). *Optimal sensor location for body sensor network to detect self-stimulatory behaviors of children with autism spectrum disorder*, United States.

Minshawi, N. F., Hurwitz, S., Fodstad, J. C., Biebl, S., Morriss, D. H., & McDougle, C. J. (2014). The association between self-injurious behaviors and autism spectrum disorders. *Psychology research and behavior management, 7*, 125-136. doi:10.2147/PRBM.S44635

Minshawi, N. F., Hurwitz, S., Morriss, D., & McDougle, C. J. (2015). Multidisciplinary assessment and treatment of self-injurious behavior in autism spectrum disorder and intellectual disability: Integration of psychological and biological theory and approach.

Journal of Autism and Developmental Disorders, 45(6), 1541-1568. doi:10.1007/s10803-014-2307-3

Murphy, O., Healy, O., & Leader, G. (2009). Risk factors for challenging behaviors among 157 children with autism spectrum disorder in ireland. *Research in Autism Spectrum Disorders, 3*(2), 474-482. doi:https://doi.org/10.1016/j.rasd.2008.09.008

Olofsen, E., W Sleigh, J., & Dahan, A. (2008). *Permutation entropy of the electroencephalogram: A measure of anaesthetic drug effect* (Vol. 101).

Ozdenizci, O., Cumpanasoiu, C., Mazefsky, C., Siegel, M., Erdogmus, D., Ioannidis, S., & Goodwin, M. S. (2018). Time-series prediction of proximal aggression onset in minimally-verbal youth with autism spectrum disorder using physiological biosignals. *arXiv preprint arXiv:1809.09948*.

Pace, G. M., Iwata, B. A., Edwards, G. L., & McCosh, K. C. (1986). Stimulus fading and transfer in the treatment of self‐restraint and self‐injurious behavior. *Journal of Applied Behavior Analysis, 19*(4), 381-389.

Pelios, L., Morren, J., Tesch, D., & Axelrod, S. (1999). The impact of functional analysis methodology on treatment choice for self‐injurious and aggressive behavior. *Journal of Applied Behavior Analysis, 32*(2), 185-195.

Plötz, T., Hammerla, N. Y., Rozga, A., Reavis, A., Call, N., & Abowd, G. D. (2012). *Automatic assessment of problem behavior in individuals with developmental disabilities.* Paper presented at the Proceedings of the 2012 ACM Conference on Ubiquitous Computing.

Poliker, R. (2006). Pattern recognition. In *Wiley Encyclopedia of Biomedical Engineering*: John Wiley & Sons, Inc.

Pugliese, C. E., Kenworthy, L., Bal, V. H., Wallace, G. L., Yerys, B. E., Maddox, B. B., . . . Anthony, L. G. (2015). Replication and comparison of the newly proposed ados-2, module 4 algorithm in asd without id: A multi-site study. *Journal of Autism and Developmental Disorders, 45*(12), 3919-3931. doi:10.1007/s10803-015-2586-3

Rad, N. M., Furlanello, C., & Kessler, F. B. (2016). Applying deep learning to stereotypical motor movement detection in autism spectrum disorders.

Rice, C. (2007). Prevalence of autism spectrum disorders--autism and developmental disabilities monitoring network, six sites, united states, 2000. *MMWR Surveill Summ, 56*(1), 1-11.

Richards, C., Oliver, C., Nelson, L., & Moss, J. (2012). Self-injurious behaviour in individuals with autism spectrum disorder and intellectual disability. *56*(5), 476-489. doi:doi:10.1111/j.1365-2788.2012.01537.x

Richardson, M. J., Schmidt, R. C., & Kay, B. A. (2007). Distinguishing the noise and attractor strength of coordinated limb movements using recurrence analysis. *Biological Cybernetics, 96*(1), 59-78. doi:10.1007/s00422-006-0104-6

Richman, J., & Moorman, J. (2000). *Physiological time-series analysis using approximate entropy and sample entropy* (Vol. 278).

Rinehart, N. J., Tonge, B. J., Bradshaw, J. L., Iansek, R., Enticott, P. G., & McGinley, J. (2006). Gait function in high-functioning autism and asperger's disorder. *European Child & Adolescent Psychiatry, 15*(5), 256-264. doi:10.1007/s00787-006-0530-y

Robertson, G., Caldwell, G., Hamill, J., Kamen, G., & Whittlesey, S. (2013). *Research methods in biomechanics, 2e*: Human Kinetics.

Robusto, K. M., & Trost, S. G. (2012). Comparison of three generations of actigraph™ activity monitors in children and adolescents. *Journal of Sports Sciences, 30*(13), 1429-1435. doi:10.1080/02640414.2012.710761

Rojahn, J., Matson, J. L., Lott, D., Esbensen, A. J., & Smalls, Y. (2001). The behavior problems inventory: An instrument for the assessment of self-injury, stereotyped behavior, and aggression/destruction in individuals with developmental disabilities. *Journal of Autism and Developmental Disorders, 31*(6), 577-588.

Rooker, G. W., Hausman, N. L., Fisher, A. B., Gregory, M. K., Lawell, J. L., & Hagopian, L. P. (2018). Classification of injuries observed in functional classes of self-injurious behaviour. *62*(12), 1086-1096. doi:doi:10.1111/jir.12535

Rosenstein, M., J. Collins, J., & De Luca, C. (1993). *A practical method for calculating largest lyapunov exponents from small data set* (Vol. 65).

Samani, A., Srinivasan, D., Mathiassen, S. E., & Madeleine, P. (2015). Nonlinear metrics assessing motor variability in a standardized pipetting task: Between-and within-subject variance components. *Journal of Electromyography and Kinesiology, 25*(3), 557-564.

Sandman, C. A., Kemp, A. S., Mabini, C., Pincus, D., & Magnusson, M. (2012). The role of self‐injury in the organisation of behaviour. *Journal of Intellectual Disability Research, 56*(5), 516-526. doi:10.1111/j.1365-2788.2012.01552.x

Schaeffer, K. M., Hamilton, K. A., & Johnson, W. L. B. (2016). Video self-modeling interventions for students with autism spectrum disorder. *Intervention in School and Clinic*, 1053451216630281.

Soares, D. A., Vannest, K. J., & Harrison, J. (2009). Computer aided self-monitoring to increase academic production and reduce self-injurious behavior in a child with autism. *Behavioral Interventions, 24*(3), 171-183. doi:10.1002/bin.283

Socie, M. J., Motl, R. W., & Sosnoff, J. J. (2014). Examination of spatiotemporal gait parameters during the 6-min walk in individuals with multiple sclerosis. *International Journal of Rehabilitation Research, 37*(4), 311-316.

Stergiou, N. (2004). *Innovative analyses of human movement*: Human Kinetics.

Tabachnick B. G. and L. S. Fidell. *Using multivariate statistics*. Boston: Pearson Education, 2013.

Tarbox, J., Wilke, A. E., Najdowski, A. C., Findel-Pyles, R. S., Balasanyan, S., Caveney, A. C., . . . Tia, B. (2009). Comparing indirect, descriptive, and experimental functional assessments of challenging behavior in children with autism. *Journal of Developmental and Physical Disabilities, 21*(6), 493-514. doi:10.1007/s10882-009-9154-8

Tibshirani, R. J. J. o. t. R. S. S. S. B. (1996). Regression shrinkage and selection via the lasso. 267-288.

Torres, E. B., & Donnellan, A. M. (2015). *Autism: The movement perspective*: Frontiers Media SA.

Torres, E. B., Yanovich, P., & Metaxas, D. N. (2013). Give spontaneity and self-discovery a chance in asd: Spontaneous peripheral limb variability as a proxy to evoke centrally driven intentional acts. *Front Integr Neurosci, 7*, 46. doi:10.3389/fnint.2013.00046

Trost, S. G., Zheng, Y., & Wong, W.-K. J. P. m. (2014). Machine learning for activity recognition: Hip versus wrist data. *35*(11), 2183.

Vohra, R., Madhavan, S., Sambamoorthi, U., & St Peter, C. (2013). Access to services, quality of care, and family impact for children with autism, other developmental disabilities, and other mental health conditions. *Autism, 18*(7), 815-826. doi:10.1177/1362361313512902

Webb, G. I. (2010). Model evaluation. In C. Sammut & G. I. Webb (Eds.), *Encyclopedia of machine learning* (pp. 683-683). Boston, MA: Springer US.

Williams, S. K., Johnson, C., & Sukhodolsky, D. G. (2005). The role of the school psychologist in the inclusive education of school-age children with autism spectrum disorders. *Journal of School Psychology, 43*(2), 117-136.

Wurdeman, S. R., Myers, S. A., & Stergiou, N. (2013). Transtibial amputee joint motion has increased attractor divergence during walking compared to non-amputee gait. *Annals of Biomedical Engineering, 41*(4), 806-813. doi:10.1007/s10439-012-0705-2

Yang, H. (2011). Multiscale recurrence quantification analysis of spatial cardiac vectorcardiogram signals. *IEEE Transactions on Biomedical Engineering, 58*(2), 339-347. doi:10.1109/TBME.2010.2063704

Yentes, J. M., Hunt, N., Schmid, K. K., Kaipust, J. P., McGrath, D., & Stergiou, N. (2013). The appropriate use of approximate entropy and sample entropy with short data sets. *Ann Biomed Eng, 41*(2), 349-365. doi:10.1007/s10439-012-0668-3

Zbilut, J. P., & Webber, C. L. (2006). Recurrence quantification analysis. In *Wiley encyclopedia of biomedical engineering*: John Wiley & Sons, Inc.

Zheng, Y.-L., Ding, X.-R., Poon, C. C. Y., Lo, B. P. L., Zhang, H., Zhou, X.-L., . . . Zhang, Y.-T. (2014). Unobtrusive sensing and wearable devices for health informatics. *IEEE Transactions on Biomedical Engineering, 61*(5), 1538-1554.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological bulletin, 99*(3), 432-442. doi:10.1037/0033-2909.99.3.432

Chapter 5: Conclusions and Recommendations

Self-injurious behavior (SIB) is among the most dangerous concerns in autism spectrum disorder (ASD). Approximately half of individuals with ASD show SIB (McTiernan, Leader, Healy, & Mannion, 2011; Murphy, Healy, & Leader, 2009; Richards, Oliver, Nelson, & Moss, 2012), which can lead to injury or hospitalization (Minshawi et al., 2014; Rooker et al., 2018). Current management methods require experience and tedious coding, and suffer from low accuracy and validity while demanding high compliance from caregivers (Marcu et al., 2013; Tarbox et al., 2009). Technological approaches could address these limitations, though such approaches have not been assessed for SIB monitoring. Novel methods could provide a "smart", automatic identification system to objectively track SIB from continuous movement data. However, ASD is a highly heterogeneous diagnosis, and SIB has a wide range of presentations, and this diversity presents a complex classification problem when implementing monitoring technology (Amaral, Schumann, & Nordahl, 2008; Bone et al., 2015; Gowen & Hamilton, 2013). Proposed approaches to this classification problem, along with recommendations for technology acceptance, are central contributions of this dissertation.

The long-term goal of this work is to develop a monitoring system capable of automatically tracking SIB and alerting caregivers and individuals of an upcoming episode. This dissertation supports the goal by examining design considerations founded on caregiver needs, as well as by evaluating potential SIB classification models and identifying effective predictors. Specifically, this research analyzed a qualitative dataset of caregiver experiences and a quantitative dataset of SIB movement to address the following aims:

- Develop guidelines for designing an SIB monitoring system from the needs of caregivers and reported needs of individuals with ASD (Chapter 2)
- Evaluate diverse classification methods to detect SIB with high accuracy and efficiency, using selected technology from established guidelines (Chapter 2) to collect movement data from individuals with ASD (Chapter 3)

- Develop a group-level model to identify a range of SIB. Examine novel predictors of SIB and implement variable intercepts and slopes to account for individual variability (Chapter 4)

Summary of major results

Caregiver perspectives

Functional assessments (FA) are traditionally used to identify the triggers that could be causing SIB (Iwata et al., 1994; Pelios, Morren, Tesch, & Axelrod, 1999; Williams, Johnson, & Sukhodolsky, 2005). However, these assessments are often tedious (Dracobly, Dozier, Briggs, & Juanico, 2018; Marcu et al., 2013), and observations of SIB may be subjective and inconsistent (Dunlap, Newton, Fox, Benito, & Vaughn, 2001; Kirby, Boyd, Williams, Faldowski, & Baranek, 2016). Recent technological innovations (i.e., sensors paired with machine learning) have the potential to address the limitations of FA. The needs of caregivers and individuals with ASD, though, should be considered before developing "smart" monitoring methods. Specifically, caregiver inclusion in interventions can encourage intervention effectiveness and support technology use for individuals with disabilities (Archer, Keshavjee, Demers, & Lee, 2014), yet caregivers and individuals with disabilities are commonly excluded from the design and evaluation of management technology (Krahn, Klein Walker, & Correa-De-Araujo, 2015; Liptak et al., 2008; Or & Karsh, 2009; Soares, Vannest, & Harrison, 2009; Vohra, Madhavan, Sambamoorthi, & St Peter, 2013).

Caregivers are critical contributors during monitoring method development, especially given their potential interactions with health technology (Pandolfe, Wright, Slack, & Safran, 2018). Designers can align new technology with current circumstances by including caregiver experiences in the design process, specifically considering the context (e.g., school or home) of the deployed technology (Valdez, Holden, Novak, & Veinot, 2014). Both educators and parents/guardians ("caregivers") of children with ASD and SIB were therefore included in interviews and focus groups to evaluate the needs pertaining to SIB and monitoring methods (Chapter 2). Because of the prevalent communication difficulties found in ASD (APA, 2013), caregivers reported the observed needs of their children/students with ASD and SIB. In total, 39 caregivers described experiences with SIB management and related technology. Qualitative content analysis was used to generate descriptive categories modified for the design community, resulting in suggestions for SIB monitoring design (Graneheim & Lundman, 2004; Hsieh & Shannon, 2005; Valdez et al., 2014).

Analyses of caregiver responses revealed seven main themes associated with SIB and technology: triggers, emotional responses, SIB characteristics, management, caregiver impact, child/student impact, and sensory/technology preferences. Resulting themes indicated areas of focus for SIB monitoring, which could extend more generally to health technology for ASD. Technology type and location preferences (e.g., transferable, easily removed yet durable, discretely placed) were particularly relevant for application in the following study (Chapter 3), and for the long-term goal of developing an accepted monitoring system to alert both caregivers and individuals with SIB of episode onset.

Evaluating machine learning for SIB detection

Monitoring systems that continuously collect data and detect SIB could support SIB management. Previous research showed the feasibility of using “smart” technology to detect other behaviors in ASD, such as stereotypical motor movements and aggression (e.g., Goodwin, Intille, Velicer, & Groden, 2008 and Ozdenizci et al., 2018). However, these studies did not specifically focus on SIB, and the application of SIB monitoring technology remains underexplored.

Sensor selection and event classification are both critical to building an automatic detection system. As initial groundwork towards determining the potential of SIB detection, we applied a wide range of supervised machine learning algorithms to SIB movement data (Chapter 3). We collected and analyzed natural movements from 10 children aged 5-14 with ASD and SIB at a child study center. The specific sensor type (wearable accelerometers) was chosen from design recommendations developed in Chapter 2. Two to seven accelerometers were placed on the wrists and waist (Min, Tewfik, Kim, & Menard, 2009), as well as on the ankles and in pockets. To examine the feasibility of automatically detecting SIB, acceleration data were then used as input into classification models developed using diverse machine learning algorithms. A wide range of methods were evaluated to detect and classify SIB, as this study is the first known application of machine learning to SIB.

High classification accuracy (up to 99.1%) was found for models trained and tested at the individual level. Both k-nearest neighbor and support vector machines outperformed other classifiers when detecting SIB, as well as when classifying into several types of SIB. A general group model, including all participants, had substantially lower classification accuracy (<65% for detection), implying that additional work is needed to create an accurate group-level model. SIB that involves finer movements (e.g., scratching) was classified with relatively strong performance metrics (i.e., specificity and F-score > 0.99), indicating the feasibility of tracking more subtle SIB with accelerometers and machine learning algorithms. Results of this study also emphasize the potential effectiveness of using wearable technology together with machine learning algorithms to detect SIB and classify types of SIB, at least for individuals.

Examining multi-level classification with nonlinear predictors of SIB

Automatic SIB identification is a complex classification problem due to highly heterogeneous ASD diagnoses and a broad range of SIB types (Amaral et al., 2008; Bone et al., 2015; Gowen & Hamilton, 2013). This inherent variability may contribute to the lower accuracy of a group-level model when compared to individual level models in prior work (Ozdenizci et al., 2018). Though individual models of movement in ASD have been reported to have accuracy up to 99%, (Albinali, Goodwin, & Intille, 2011; Coronato, De Pietro, & Paragliola, 2014; Goodwin et al., 2014; Rad, Furlanello, & Kessler, 2016), they may not generalize beyond specific participant datasets. Group-level models, in contrast, could provide the efficiency needed for environments in which there is more than one child with SIB (e.g., classrooms or a clinics). Further, nonlinear motor variability metrics that capture the aforementioned complexity in ASD and SIB (Großekathöfer et al., 2017) could serve as useful predictors of SIB and contribute to enhanced group-level accuracy in SIB detection.

In an effort to create an efficient and more accurate group-level classifier, SIB data from Chapter 3 were used as input into a multi-level logistic regression model. Time-domain, frequency-domain, and nonlinear features were created, and lasso and principal component analysis were used for feature selection and reduction. Variable slopes and intercepts were included in the regression model to allow for individual differences within the group. This multi-level model with novel features was compared to several other diverse classifiers, such as stepwise regression, machine learning algorithms (e.g., kNN and SVM), and individually trained models.

The multi-level logistic regression model, with variable intercepts and slopes, identified SIB with ~75% accuracy among a group of heterogeneous participants and behaviors. This accuracy level was superior (by ~6%) to previous work that similarly examined ASD movement in a group of participants (Ozdenizci et al., 2018). Nonlinear motor variability metrics (e.g., recurrence quantification analysis metrics) contributed as significant predictors of SIB, suggesting the presence of nonlinear variability in the patterns of SIB. This research was the first to our knowledge to examine both a multi-level group model of SIB and nonlinear properties in the movements of SIB.

Suggestions for future research

The current work has demonstrated the feasibility of developing an SIB monitoring system to detect SIB with high accuracy and efficiency. Design considerations derived from caregiver needs (Chapter 2) guided technology selection to collect movement data, while diverse classification methods provided groundwork for effective SIB modeling (Chapters 3 and 4). This dissertation serves as an initial effort towards automatic SIB tracking, though research should continue to enhance classification methods and work toward an alert system for SIB management. Specific areas of potential future work include directly incorporating child experiences to influence system development, exploring additional advanced modeling techniques, further analyzing SIB through a dynamical systems perspective, and designing a management interface for caregivers and children with ASD.

Though our work adopted a user-centered perspective as related to caregivers (Chapter 2), the nonverbal or minimally verbal characteristics of the participant pool precluded this same approach for children with ASD and SIB. Children with ASD and SIB in general are likely to be minimally or nonverbal (Richards, Moss, Nelson, & Oliver, 2016), which can be a barrier to collecting their personal experiences. Child perspectives were considered here through observation, and two participants who communicated their thoughts about wearable sensors (Chapter 2), yet this effort should be expanded upon in future work. For example, children with communicative abilities (whether verbal or nonverbal) and SIB could be recruited specifically in a usability study. Children could be given different options for sensors, to select which sensor they would use, as well as which location supports their use and if there are preferred materials. Reactions could be monitored across environments, such as in the classroom, where social interactions can affect technology preferences (Giesbrecht, 2013). Such additional work reflects user-centered design, and could thereby support management effectiveness by encouraging compliance with SIB tracking.

Sophisticated additions to the classification models presented here could address both the evolving nature of SIB and the substantial inherent individual variability. Such additions include semi-supervised methods and the use of Bayesian priors. Supervised methods were examined in this work as a first effort toward classifying SIB (Chapter 3), yet semi-supervised approaches could provide the adaptability needed to detect new behaviors (Jain, Duin, & Mao, 2000). The first study (Chapter 2) revealed that SIB can change forms completely, for example from head banging to self-biting. A semi-supervised approach would allow initial training on the present SIB, while offering the capability of learning future SIB. Information specific to an individual could also be addressed through including Bayesian priors. Coefficient estimates can also update with each new episode of SIB, using an "adaptive" empirical Bayes approach. Prior distributions can be computed from the mean coefficients of the training data, and the posterior distribution can be recomputed with data in real-time. Continuously updating predictors might help to improve current models, and aid in addressing evolving behavior when tracking SIB.

In addition to the model advancements described above, prediction techniques can be used to alert caregivers of an oncoming SIB episode. For example, Bayesian networks represent multivariate probability distributions, and provide a conditional probability of the event of interest (Sammut & Webb, 2010). This probability could be compared to a determined threshold created from caregiver preferences, with risk based on behavioral severity, behavior escalation trends, and SIB frequency tendencies specific to that child. For example, caregivers could be alerted if the probability exceeds 0.5 for one child having severe SIB with sudden onset, but alerted if the probability exceeds 0.9 for another child who shows more subtle and infrequent SIB with gradual onset. SIB prediction can also be explored through training classifiers with labels of “1” at a determined time period *before* SIB, and “0” otherwise. If a classifier can detect an SIB episode before it occurs, the model would then have a predictive capability. The ability to predict an episode before its occurrence is supported by previous work, which implemented a regression model to predict aggression towards others up to 60 seconds before occurrence (Ozdenizci et al., 2018). Future research should therefore apply predictive techniques to extend the models explored in this work.

The current work is the first known to apply a dynamical movement perspective to SIB (Chapter 4). Predictors capturing nonlinear aspects of movement contributed significantly to the regression model, indicating nonlinear trends in the movements of SIB. This finding should be analyzed further, such as by comparing nonlinear metrics of motor variability between SIB versus non-SIB movements. Such comparisons could provide information about underlying mechanisms of SIB, possibly supporting theories that SIB functions to organize behavioral patterns and establish certain levels of motor complexity (Sandman, Kemp, Mabini, Pincus, & Magnusson, 2012). Also, nonlinear variables derived for feature selection and reduction were chosen as a starting point for examining nonlinear predictors of SIB, based on suggestions from previous work that indicated their reliability for human data and short time windows (Großekathöfer et al., 2017; Samani, Srinivasan, Mathiassen, & Madeleine, 2015; Yentes et al., 2013). However, other variables should also be examined in future studies, such as different types of entropy (Yentes et al., 2013) or the Lyapunov exponent (Wolf, Swift, Swinney, & Vastano, 1985). Additional nonlinear analyses could provide a more complete picture of the motor characteristics of SIB.

The described extensions of the current SIB identification methods could lead to an SIB monitoring system to alert caregivers of ensuing episodes. Future work will likely need to focus on interface design to create an effective platform for clinicians, caregivers, and individuals with ASD to view and interpret data. Input options could include management methods suggested by the clinician, and after episodes occur, caregivers could use minimal input (e.g., selecting options on buttons) to note the methods used and results of those methods (e.g., stopping SIB). Data analyses could further lead to management suggestions. For example, when SIB is detected, a caregiver could receive a phone notification that, “Redirecting to squeezing a stress ball is 80% effective, stopping SIB within 20 seconds,” to help direct management decisions. Though the design of monitoring technology was extensively discussed with caregivers in this work (Chapter 2), caregiver and child needs related to interface design should also be examined. In particular, the sensory preferences that often to present in children with ASD should be considered when designing the visual, auditory, and/or haptic feedback of an interface. Effective feedback (e.g., vibration) on an interface could orient the child to a predefined activity at the first sign of SIB onset, promoting consistency in management. Thus, in conjunction with advanced sensing technology and modeling, an effective interface would place SIB monitoring and management into the hands of caregivers and individuals with SIB.

Overall conclusions

This research offers design guidelines and presents individual- and group-level models for an SIB monitoring system. The work supports the feasibility of using wearable technology and classification methods to accurately and efficiently identify SIB and to classify distinct types of SIB. Additional research is needed, though, to support continued development of a tracking system, specifically through exploring adaptive methods for evolving behaviors. Future work is also recommended that applies SIB identification methods to create an alert system for prompt SIB management from caregivers, and for self-awareness interventions for those individuals with ASD.

References

Albinali, F., Goodwin, M. S., & Intille, S. (2011). Detecting stereotypical motor movements in the classroom using accelerometry and pattern recognition algorithms. *Pervasive and Mobile Computing, 8*(1), 103. doi:10.1016/j.pmcj.2011.04.006

Amaral, D. G., Schumann, C. M., & Nordahl, C. W. (2008). Neuroanatomy of autism. *Trends in Neurosciences, 31*(3), 137-145. doi:https://doi.org/10.1016/j.tins.2007.12.005

APA. (2013). *Dsm 5*: American Psychiatric Association.

Archer, N., Keshavjee, K., Demers, C., & Lee, R. (2014). Online self-management interventions for chronically ill patients: Cognitive impairment and technology issues. *INTERNATIONAL JOURNAL OF MEDICAL INFORMATICS, 83*(4), 264-272. doi:https://doi.org/10.1016/j.ijmedinf.2014.01.005

Bone, D., Goodwin, M. S., Black, M. P., Lee, C.-C., Audhkhasi, K., & Narayanan, S. (2015). Applying machine learning to facilitate autism diagnostics: Pitfalls and promises. *Journal of Autism and Developmental Disorders, 45*(5), 1121-1136. doi:10.1007/s10803-014-2268-6

Coronato, A., De Pietro, G., & Paragliola, G. (2014). A situation-aware system for the detection of motion disorders of patients with autism spectrum disorders. *Expert Systems with Applications, 41*(17), 7868-7877.

Dracobly, J. D., Dozier, C. L., Briggs, A. M., & Juanico, J. F. (2018). Reliability and validity of indirect assessment outcomes: Experts versus caregivers. *Learning and Motivation, 62*, 77-90. doi:https://doi.org/10.1016/j.lmot.2017.02.007

Dunlap, G., Newton, J. S., Fox, L., Benito, N., & Vaughn, B. (2001). Family involvement in functional assessment and positive behavior support. *Focus on autism and other developmental disabilities, 16*(4), 215-221.

Giesbrecht, E. (2013). Application of the human activity assistive technology model for occupational therapy research. *Australian Occupational Therapy Journal, 60*(4), 230-240.

Goodwin, M., Haghighi, M., Tang, Q., Akcakaya, M., Erdogmus, D., & Intille, S. (2014, 2014). *Moving towards a real-time system for automatically recognizing stereotypical motor movements in individuals on the autism spectrum using wireless accelerometry*.

Goodwin, M. S., Intille, S. S., Velicer, W. F., & Groden, J. (2008). *Sensor-enabled detection of stereotypical motor movements in persons with autism spectrum disorder.* Paper presented at the Proceedings of the 7th international conference on Interaction design and children.

Gowen, E., & Hamilton, A. (2013). Motor abilities in autism: A review using a computational context. *Journal of Autism and Developmental Disorders, 43*(2), 323. doi:10.1007/sl0803-012-1574-0

Graneheim, U. H., & Lundman, B. (2004). Qualitative content analysis in nursing research: Concepts, procedures and measures to achieve trustworthiness. *Nurse education today, 24*(2), 105-112.

Großekathöfer, U., Manyakov, N. V., Mihajlović, V., Pandina, G., Skalkin, A., Ness, S., . . . Goodwin, M. S. (2017). Automated detection of stereotypical motor movements in autism spectrum disorder using recurrence quantification analysis. *Frontiers in Neuroinformatics, 11*, 9. doi:10.3389/fninf.2017.00009

Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative health research, 15*(9), 1277-1288.

Iwata, B. A., Pace, G. M., Dorsey, M. F., Zarcone, J. R., Vollmer, T. R., Smith, R. G., . . . Willis, K. D. (1994). The functions of self-injurious behavior: An experimental-epidemiological analysis. *27*(2), 215-240. doi:doi:10.1901/jaba.1994.27-215

Jain, A. K., Duin, R. P. W., & Mao, J. (2000). Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22*(1), 4-37. doi:10.1109/34.824819

Kirby, A. V., Boyd, B. A., Williams, K. L., Faldowski, R. A., & Baranek, G. T. (2016). Sensory and repetitive behaviors among children with autism spectrum disorder at home. *Autism*, 1-13.

Krahn, G. L., Klein Walker, D., & Correa-De-Araujo, R. (2015). Persons with disabilities as an unrecognized health disparity population. *American Journal of Public Health, 105*(S2), S198-S206. doi:10.2105/AJPH.2014.302182

Liptak, G. S., Benzoni, L. B., Mruzek, D. W., Nolan, K. W., Thingvoll, M. A., Wade, C. M., & Fryer, G. E. (2008). Disparities in diagnosis and access to health services for children with autism: Data from the national survey of children's health. *J Dev Behav Pediatr, 29*(3), 152-160. doi:10.1097/DBP.0b013e318165c7a0

Marcu, G., Tassini, K., Carlson, Q., Goodwyn, J., Rivkin, G., Schaefer, K. J., . . . Kiesler, S. (2013). *Why do they still use paper?: Understanding data collection and use in autism education*. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Paris, France.

McTiernan, A., Leader, G., Healy, O., & Mannion, A. (2011). Analysis of risk factors and early predictors of challenging behavior for children with autism spectrum disorder. *Research in Autism Spectrum Disorders, 5*(3), 1215-1222. doi:https://doi.org/10.1016/j.rasd.2011.01.009

Min, C.-H., Tewfik, A. H., Kim, Y., & Menard, R. (2009, 2009). *Optimal sensor location for body sensor network to detect self-stimulatory behaviors of children with autism spectrum disorder*, United States.

Minshawi, N. F., Hurwitz, S., Fodstad, J. C., Biebl, S., Morriss, D. H., & McDougle, C. J. (2014). The association between self-injurious behaviors and autism spectrum disorders. *Psychology research and behavior management, 7*, 125-136. doi:10.2147/PRBM.S44635

Murphy, O., Healy, O., & Leader, G. (2009). Risk factors for challenging behaviors among 157 children with autism spectrum disorder in ireland. *Research in Autism Spectrum Disorders, 3*(2), 474-482. doi:https://doi.org/10.1016/j.rasd.2008.09.008

Or, C. K., & Karsh, B. T. (2009). A systematic review of patient acceptance of consumer health information technology. *J Am Med Inform Assoc, 16*(4), 550-560. doi:10.1197/jamia.M2888

Ozdenizci, O., Cumpanasoiu, C., Mazefsky, C., Siegel, M., Erdogmus, D., Ioannidis, S., & Goodwin, M. S. (2018). Time-series prediction of proximal aggression onset in minimally-verbal youth with autism spectrum disorder using physiological biosignals. *arXiv preprint arXiv:1809.09948*.

Pandolfe, F., Wright, A., Slack, W. V., & Safran, C. (2018). Rethinking the outpatient medication list: Increasing patient activation and education while architecting for centralization and improved medication reconciliation. *Journal of the American Medical Informatics Association, 25*(8), 1047-1053. doi:10.1093/jamia/ocy047

Pelios, L., Morren, J., Tesch, D., & Axelrod, S. (1999). The impact of functional analysis methodology on treatment choice for self‐injurious and aggressive behavior. *Journal of Applied Behavior Analysis, 32*(2), 185-195.

Rad, N. M., Furlanello, C., & Kessler, F. B. (2016). Applying deep learning to stereotypical motor movement detection in autism spectrum disorders.

Richards, C., Moss, J., Nelson, L., & Oliver, C. (2016). Persistence of self-injurious behaviour in autism spectrum disorder over 3 years: A prospective cohort study of risk markers. *J Neurodev Disord, 8*, 21. doi:10.1186/s11689-016-9153-x

Richards, C., Oliver, C., Nelson, L., & Moss, J. (2012). Self-injurious behaviour in individuals with autism spectrum disorder and intellectual disability. *56*(5), 476-489. doi:doi:10.1111/j.1365-2788.2012.01537.x

Rooker, G. W., Hausman, N. L., Fisher, A. B., Gregory, M. K., Lawell, J. L., & Hagopian, L. P. (2018). Classification of injuries observed in functional classes of self-injurious behaviour. *62*(12), 1086-1096. doi:doi:10.1111/jir.12535

Samani, A., Srinivasan, D., Mathiassen, S. E., & Madeleine, P. (2015). Nonlinear metrics assessing motor variability in a standardized pipetting task: Between-and within-subject variance components. *Journal of Electromyography and Kinesiology, 25*(3), 557-564.

Sammut, C., & Webb, G. I. (2010). Bayesian network. In. (Ed.), *Encyclopedia of machine learning* (pp. 81-81). Boston, MA: Springer US.

Sandman, C. A., Kemp, A. S., Mabini, C., Pincus, D., & Magnusson, M. (2012). The role of self‐injury in the organisation of behaviour. *Journal of Intellectual Disability Research, 56*(5), 516-526. doi:10.1111/j.1365-2788.2012.01552.x

Soares, D. A., Vannest, K. J., & Harrison, J. (2009). Computer aided self-monitoring to increase academic production and reduce self-injurious behavior in a child with autism. *Behavioral Interventions, 24*(3), 171-183. doi:10.1002/bin.283

Tarbox, J., Wilke, A. E., Najdowski, A. C., Findel-Pyles, R. S., Balasanyan, S., Caveney, A. C., . . . Tia, B. (2009). Comparing indirect, descriptive, and experimental functional assessments of challenging behavior in children with autism. *Journal of Developmental and Physical Disabilities, 21*(6), 493-514. doi:10.1007/s10882-009-9154-8

Valdez, R. S., Holden, R. J., Novak, L. L., & Veinot, T. C. (2014). Transforming consumer health informatics through a patient work framework: Connecting patients to context. *Journal of the American Medical Informatics Association, 22*(1), 2-10.

Vohra, R., Madhavan, S., Sambamoorthi, U., & St Peter, C. (2013). Access to services, quality of care, and family impact for children with autism, other developmental disabilities, and other mental health conditions. *Autism, 18*(7), 815-826. doi:10.1177/1362361313512902

Williams, S. K., Johnson, C., & Sukhodolsky, D. G. (2005). The role of the school psychologist in the inclusive education of school-age children with autism spectrum disorders. *Journal of School Psychology, 43*(2), 117-136.

Wolf, A., Swift, J. B., Swinney, H. L., & Vastano, J. A. (1985). Determining lyapunov exponents from a time series. *Physica D: Nonlinear Phenomena, 16*(3), 285-317. doi:http://dx.doi.org/10.1016/0167-2789(85)90011-9

Yentes, J. M., Hunt, N., Schmid, K. K., Kaipust, J. P., McGrath, D., & Stergiou, N. (2013). The appropriate use of approximate entropy and sample entropy with short data sets. *Ann Biomed Eng, 41*(2), 349-365. doi:10.1007/s10439-012-0668-3

Appendices

Appendix A – Chapter 2

Themes and design considerations

Table A1. Themes and subcategories derived from parent and educator responses with supporting participant quotes underlying the results. (P) = parent quote, (E) = educator quote

Themes	Subcategories	Participant quotes
Triggers	• Cognitive Rigidity (difficulty changing mental states) • Demands (e.g., school assignment) • Fatigue • Communication • Medical Factors and Health • Relationships • Sensory Stimulation (under or overwhelmed) • Situational Variability (task/setting transitions) • Unmet Need (e.g., thirst, attention)	• "As far as when we're with him, I think it will happen just maybe when he's tired and there's been a lot of going on and people in and out" (P) • "I think a lot of kids also really like interpersonal contact, but are unable to express that that's what they want..." (E)
Emotional Responses	• Angry • Excited • Frustrated • Stressed • Upset	• "It's almost like he had to get some sort of enjoyment out of it"(P) • "I would say just the biggest factor would be stress and then, change in environment or a change in routine or change in schedule, which leads to stress" (E)

Self-Injurious Behavior Characteristics	• Age of Onset • Change over Time • Presentation • Duration • Frequency • Intensity • Location on Body • Position of Body • Setting • Time • Environment Use • Object Use • Group Trends • Physical Result (often discussed in terms of intensity)	• "Oh it was probably before we even had a diagnosis, probably around two-and-a half" (P) • "He takes his head and hits it against a wall, a window" (P) • "We've seen students leave for a holiday…they leave with 7,000 SIB's in six hours and they come back and they have 700 for months." (E)
Management Approaches	• Resources • Coordination (collaboration, caregiver consistency) • Sense-making (data collection, often informing strategies if completed) • Strategies (e.g., ignoring, redirection, token systems) • Effectiveness	• "It was when the structure started to not be so structured is when you would see he really needs the support. He does. He needs to be constantly praised, and constantly reminded. He thrives that way." (P) • "The best approach is to do it 24/7, and everybody has to be on the same page, including us."(P) • "So we have this huge double-layer sock filled with rice so he can do that [squeeze] to that instead of his arm..." (E) • "Then by the time it stops or you've got that child calmed down, you're forgetting,

		'Oh, I've got to do that [take data].'" (P) • "I also can't hold a blocking pad and click how many times I just got punched in the face" (E)
Child/Student Impact	• Educational • Emotional • Medical (e.g., tissue damage, loss of eyesight, concussions) • Relational	• "Even the social part for her, wanting friends, but yet neuro-typical kids realize that something's different, and they don't want be her friend, …or they will abuse the friendship just to see her get in trouble, or do something that they think is funny. It's not funny." (P) • "The biggest concern is [students] not being able to access the same type of educational setting or program that they would otherwise be able to access because of such extreme behavior." (E)
Caregiver Impact	• Financial • Emotional • Medical • Relational • Occupational	• "[I] wept a lot because I feel so helpless" (P) • "The [school] cost is more than what my husband makes in a year." (P) • "Until you've lived this life, you really don't understand the emotional battle that comes with it, that can cause issues between a wife, and a husband, grandparents,

		siblings, neighbors. The emotional part is as hard as the physical part, sometimes." (P) • "Also, just seeing a person, you see a 120-pound little girl bleeding from every place on her face, and your instructions are not to reinforce that. That itself is just tough." (E)
Sensory/Technology Preferences *Note: not specific to SIB*	• Visual • Auditory (e.g., repeated movie lines) • Olfactory (e.g., scented candles) • Tactile (e.g., sequins, soft material) • Gustation/Mouthing (e.g., chewing) • "Touch+"/Haptic Feedback (e.g., physical pressure or vibration) • Technology Use (present and potential)	• "It would take him about half a day to get used to a hospital band on his ankle, and then finally he'll leave it alone…it's sensory" (P)" • "Put it [sensor] on their back where they can't reach it." (P) • "Well, the timer, he does use the timer. He will rely on that. If he is given a 'nice hands,' he does watch that one-minute count down. He does like to look at the clock, because his schedule is pretty rigid. He knows. He will check the clock. He will check on lunch time." (P) • "Yeah, everything's got to be breakaway. I mean, in the context of technology, putting anything worth more than $10.00 that's breakaway, probably not going to be the best for our students." (E) • "I just can't believe that

		we're in 2017 and we're all taking [SIB data] on paper. We're all writing it all down. Then, we go into the computer and enter it into Excel" (E)

Table A2. Design considerations for SIB prediction technology.

Themes	Descriptive Guidance	Prescriptive Guidance
Triggers	• Monitor external situations that may trigger the behavior • Consider data that may relate to trigger occurrence • Establish a unified platform to integrate all stakeholder input on different triggers • Establish a method to relay trigger suggestions to caregivers	• Equip child/student, caregivers or environment with technology measuring physiological signals, patterns of movement, and/or environmental feedback like location services, noise or light to examine potential triggers • Provide user input and machine self-learning capabilities to customize triggers, such as a noise level threshold, to be detected from the environment • Accept and store user input from select caregivers (e.g., educators and parents) • Allow secure, cloud-based storage of user input and data collected from technology for select caregiver access through a password-protected portal • Factor caregiver input of SIB observations into machine learning algorithms for SIB detection and prediction • Continuously collect and mine data to determine patterns of detected SIB events

		• Offer an accessible user interface, such as a smart phone application, for the caregiver to interact with the system and receive notifications
Emotional Responses	• Account for emotions that may precede or coincide with SIB • Include positive (e.g., excitement), negative (e.g., anger) and neutral responses	• Collect physiological signals (Electrodermal activity, heart rate, temperature), movements (acceleration, motion tracking) and/or noise levels that may relate to emotional response • Use combinations of sensor data to distinguish emotions with similar physiological signals (e.g., anger and excitement) that may lead to different behaviors • Offer opportunities for caregiver input to assist in emotion recognition (e.g., scales) for data mining • Provide input options for child/student to select emotions he/she feels (e.g., emoticons) • Include feedback like vibrations or a visual notification to encourage child/student to assess his/her emotions

Self-Injurious Behavior Characteristics	• Account for age/changes with age • Provide a robust method for a variety of potentially evolving behaviors • Provide flexible monitoring methods to track different parts of the body • Provide adaptable and transferrable methods for different environments and times of day • Establish durable sensors to tolerate aggression toward objects • Avoid hazardous locations in environment or on the body that may be used for SIB	• Develop multi-level algorithms for different age ranges or other personal characteristics • Include adaptive predictors within machine learning algorithms to adjust according to varying characteristics, like behavior type or sensor placement • Measure duration and frequency of behaviors, such as through accelerometry • Consider objective measures of intensity, like change in acceleration over time • Use sensor fusion techniques to accommodate multiple locations where SIB is presented, like wearable devices for use outside the home/school and nonwearable devices in homes/classrooms
Management Approaches	• Provide means for caregivers to acquire management support	• Allow all clinicians of the child/student to have access to an unified portal, like a

	• Offer accessible data on behavioral changes for all collaborating caregivers • Provide remote capabilities for standardized data taking and/or management methods to ensure consistency • Incorporate other management needs	website or smartphone application, for information exchange among caregivers and professionals • Display information to specified caregivers about SIB, both child/student-specific and general clinical advice, through an interactive and centralized application • Offer remote monitoring for parent/guardian to collect data and ensure consistent management from afar • Include consistent input options for caregivers (e.g., timer, counter, intensity scales) if manually tracking • Show correlation of sensor stream data and manual data • Display data visualization using integrated tracking across caregiver input and data streams • Offer combined tracking and management options such as picture communication when algorithms indicate the child is frustrated, or schedule reminders for child if he/she shows SIB during transitions
Child/Student Impact	• Supplement short-staffed schools and promote access to education • Provide hospital connection for medical emergencies	• Automate methods for data collection, SIB prediction and management administration to

	• Consider socially acceptable technology	support education options for the student • Automatically notify a caregiver, educator, clinician or emergency staff when intensity surpasses a set threshold • Incorporate classification methods with commonly used technology, like cellphones or smart watches • Use discrete locations for technology, like pockets or under clothing
Caregiver Impact	• Consider low-cost solutions • Provide support for caregivers to access information and suggestions about SIB and its management, specific for his/her child • Provide means of educating community, friends and family about ASD/SIB • Consider opportunities for networking and information sharing among caregivers	• Incorporate devices already adopted in schools and homes, like cellphones or tablets to mitigate costs • Require a limited number of quick interactions to accommodate tight schedules • Mitigate required input in the presence of high pre-existing stress levels • Establish a cloud-

		based platform for caregivers to share experiences and information to family members and friends
Sensory/Technology Preferences	• Include customizable attachments for different tactile needs and clothing • Integrate preferred sensory stimuli as options for behavior substitution • Use tolerated or preferred material for wearable sensors, yet ensure the sensor is not promoting an undesired fixation • Consider sensory feedback to which the child/student responds, but neither instigates further SIB nor disrupts the functioning of others • Ensure the device is durable for potential sensory uses • Use secure means to connect sensing technology to the child or environment • Include easy removal strategy for caregivers for emergencies • Explore existing options of preferred technology for natural integration into everyday life	• Offer varying material coverage for wearable technology, like sequin or fleece options and difference material designs or prints • Incorporate customizable sound or light displays, as well as textures or scents • Offer safe and chewable material attachments for biting • Add preferred feedback, like vibration or movement for reward or positive notification • Add neutral feedback, like a specific color of light, for SIB onset warnings to the child or caregiver • If using wearable technology, add clasps with “break-away” options for school safety requirements that can

		be quickly removed by caregivers, yet not their children/students • Ensure wearable sensors are waterproof (i.e., washing hands or if thrown in toilets) • Ensure the device is durable for high impact (e.g., throwing, stomping, biting, repeated contact) by encasing it • Include everyday technology like timers or video to encourage use

Demographic forms

Demographic Questionnaire for Parents

Please fill out the following form:

1) Your age:
____

2) Your gender:
____

3) Number of children living in your home (total):
____

4) Do all of your children live with you full time?
___Yes
___No (please specify): ____________

5) Number of children with autism spectrum disorder (ASD) living in your home:
____

6) Number of children with ASD **AND** self-injurious behavior (for example, self-hitting or head banging) living in your home:
____

7) Please rate the degree to which self-injurious behavior is a problem for your child (circle one):

0	**1**	**2**	**3**
It is **not** a problem	It is a **minor** problem	It is a **significant** problem	It is a **severe** problem

8) Your child's age:
____

9) Your child's gender (child with self-injurious behavior):
____

10) Your relationship to the child (please select one):

___Biological parent

___Adoptive parent
___Other (please specify): ___________

11) Your relationship to the child (please select one):

___Father
___Mother
___Other (please specify): ___________

12) Your child's grade (for example, 4^{th} grade):
____

13) Your race (please select all that apply):

___American Indian or Alaskan Native
___Asian
___Black or African American
___Native Hawaiian or Other Pacific Islander
___White
___Other (please specify): ___________
___Do not wish to provide

14) Your ethnicity:

___Hispanic/Latino
___Not Hispanic/Latino
___Do not wish to provide

15) Your child's race (please select all that apply):

___American Indian or Alaskan Native
___Asian
___Black or African American
___Native Hawaiian or Other Pacific Islander
___White
___Other (please specify): ___________
___Do not wish to provide

16) Your child's ethnicity:

___Hispanic/Latino
___Not Hispanic/Latino
___Do not wish to provide

17) Highest personal education level (for example, high school):

___Less than 8^{th} grade
___Some high school
___Finished high school (or equivalent)
___Some college or AA degree or technical school
___Bachelor's degree (BA, BS)
___Post-graduate degree
___Do not wish to provide

18) Total household income:

___less than $20,000
___$21,000 to $35,000
___$36,000 to $50,000
___$51,000 to $65,000
___$66,000 to $80,000
___$81,000 to $100,000
___$101,000 to $130,000
___$131,000 to $160,000
___Over $160,000
___Do not wish to provide

Demographic Questionnaire for Educators

Please fill out the following form:

1. Your age:
   a. ___________

2. Your gender:
   a. ___________

3. Your race (please select all that apply):

   a. ___American Indian or Alaskan Native
   b. ___Asian
   c. ___Black or African American
   d. ___Native Hawaiian or Other Pacific Islander
   e. ___White
   f. ___Other (please specify): ___________
   g. __Do not wish to provide

4. Your ethnicity:

   a. ___Hispanic/Latino
   b. ___Not Hispanic/Latino
   c. __Do not wish to provide

5. Highest personal education level:

   a. __Less than 8th grade
   b. __Some high school
   c. __Finished high school (or equivalent)
   d. __Some college or AA degree or technical school
   e. __Bachelor's degree (BA, BS)
   f. __Post-graduate degree
   g. __Do not wish to provide

6. Your job title as an educator (for example, fourth grade teacher):
   a. ___________

7. Your location of work (for example, private or public school, exclusively ASD education)
   a. ________

8. Age of children with ASD that you work with:

a. ____________

9. Years of experience working with ASD (including current year of work):

    a. ____________

10. Past job titles related to your work with ASD:

    a. ____
    b. ____

Interview/focus group questions

Interview/Focus Group Questions (Parents)

1. Sometimes people with autism may use actions that look similar to self-injurious behavior. This behavior may hurt, but may not cause harm. For example, someone may hit their chest once when they are about to receive a prize. What are behaviors you have seen that look like self-injurious behaviors, but may not be? How are these behaviors different?

    a. How is your child's response to excitement different than to frustration?

2. Please describe self-injurious behavior you witness or have witnessed your child showing.

    a. When does your child usually show this behavior? (What do you think is the trigger?)
        i. Please describe examples of trends or patterns you notice about these behaviors.

    b. Where does your child usually show this behavior?
        i. Location/environment?
        ii. Parts of the body involved in this injurious behavior?

    c. *[If need prompt]* Please tell me more about these behaviors, for example, can you tell me about the frequency, intensity, and duration?

3. Tell me about the first time you noticed your child's self-injurious behavior.

    a. Please tell me if/how this behavior has changed over time. (Why do you think this change occurred)?

4. Please describe how you respond to your child's self-injurious behavior.
    a. When this behavior occurs, what are you thinking of at the moment? For example, do you first think of 1) minimization of harm, 2) stopping that moment, or 3) working towards reducing the behavior in the long-term (or something else)?

    b. After talking to parents in previous interviews, we heard that some found certain objects or technology to be helpful when trying to manage behaviors. For example, a toothbrush was offered to a girl who plucked her hair, and a favorite DVD was offered to a boy when he stopped banging his head. What do you use to manage [*or substitute other word they use*] this behavior?
        i. Are there objects you use to manage [*or substitute other word they use*] the behavior? (Further prompting: please describe any objects/technology you have *seen* used).
            1. Everyday/non-digital objects (ball, brush, blanket)?

        2. Digital technology (phones, computers)?
        3. Entertainment (video, DVD)?

5. In what areas do you feel confident when managing self-injurious behavior?

6. What do you need help with when managing self-injurious behavior?

7. What management approaches have been suggested to you?

    a. How do you feel about these suggested approaches?

    b. Have you tried any of these approaches? Why or why not?
        i. What about these worked well?
        (For further prompting: Please describe examples of approaches that helped decrease self-injurious behavior.)
        ii. What about these approaches could have been better?
    c. In what ways did you track/log changes in behavior?
        i. What makes it challenging to track/log these behaviors?
        ii. What makes it easy to track/log these behaviors?

8. Please describe the availability of support services around you for parents/guardians/caregivers of children with autism. [*follow with* "For SIB?"]

9. When do you/did you work with the school or teachers to reduce self-injurious behavior?
    a. How is the home different from the school environment when managing SIB?
    b. What are challenges in transitioning school interventions to other environments (like the home)?
    c. [If caregivers in the home, also ask, "What are difficulties of working with people who help care for your child? What are benefits?"]

10. What effect does self-injurious behavior have on your home?

11. If applicable, how do your other children respond to your child's self-injurious behavior?

12. How do other family members respond to your child's self-injurious behavior?

13. How do other people (outside of your family) respond to your child's self-injurious behavior?

14. Has the presence of self-injurious behavior changed your relationships with those around you? If so, how?

15. What other challenges come from this self-injurious behavior?

16. Please describe technology (like digital devices) [*if need example= DVD, computer, iPhone, iPad*] that you have used with your child or you have seen your child use.

a. Are any used in interventions to manage self-injurious behavior?
b. How about any objects [*if need example = pillows, bouncy balls*]?

17. How does your child respond to noise feedback, like an alarm?
    a. How about visual feedback?
        i. Touch?
        ii. Vibration?
        iii. *For each sense, if unclear:* Are these responses positive (approach the sound/image etc.) or negative (retreat from the sound/image etc.)?

18. Can you tell me how your child reacts to touch (*from you, family, self, objects)?*
    a. Are there certain parts of the body that he/she does/does not want something (clothes etc.) to touch him/her? If so, please describe.

19. What specific likes/preferences does your child have? (*If they need further clarification* - Are there particular things – objects, activities, or places – your child is drawn to or enjoys)?

    a. What are specific dislikes/aversions does your child have? (Do any specific objects/environments seem to cause self-injurious behavior or a negative reaction?)

20. How do you feel about using a technology that would not be worn, and instead would be installed in an environment or the home to track SIB automatically?
    a. What would get in the way of applying something that is in the environment/room to track SIB?

21. How do you feel about asking your child to wear something that would track SIB automatically?
    a. What would get in the way of applying something that could be worn to track SIB?
    b. [*If need prompt*] How would your child react to wearing something added to his/her body that could automatically log behavior?
        i. Smart watch/bracelet?
        ii. Hat or headband?
        iii. Necklace?
        iv. Scarf?
        v. A sleeve/cuff?
        vi. Anklet?
        vii. Socks or shoes?
        viii. Belt?
        ix. Pants?
        x. Shirt?
        xi. Any other articles of clothing or accessories that your child does or does not like?

[*If in person, show technology, discuss how it works, let people touch the technology; otherwise, send picture of technology with size, videos of holding technology, etc. if able to through email*]

[Repeat questions 19, 20; add, "What do you like about this technology?" "What would be helpful about this technology?" "What do you not like?" "What would be challenging?" "When would [type of technology] be useful/not useful?"]

22. If a team of people were trying to create methods that could help self-injurious behavior in some way, what would you want them to know?
    a. What would you want them to know [about you, your home, your family, your child] if they are trying to create something wearable/that can be worn to track this behavior?
    b. What would you want them to know [about you, your home, your family, your child] if they are trying to create something that is not wearable – like something in the room/environment- to track this behavior?

Interview/Focus Group Questions (Educators)

1. There is a lot of variability in what an autism spectrum diagnosis looks like. Please describe the range of how this diagnosis presents in the students you work with.

2. Sometimes people with autism may use actions that look similar to self-injurious behavior. This behavior may hurt, but may not cause harm. For example, someone may hit their chest once when they are about to receive a prize. What are behaviors you have seen that look like self-injurious behaviors, but may not be? How are these behaviors different?

3. Please describe different types of self-injurious behavior you have witnessed your students showing.

    a. Please describe examples of trends or patterns you notice about these behaviors
    b. *[If need prompt]* Please tell me more about these behaviors, for example, can you tell me about the frequency, intensity, and duration?

4. Please describe how you have managed students' self-injurious behaviors.
    a. When self-injurious behaviors occur, what are you thinking of at the moment? For example, do you first think of 1) minimization of harm, 2) stopping that moment, or 3) working towards reducing the behavior in the long-term (or something else)?
    b. What role does student-to-educator trust play in SIB management?
    c. After talking to educators in previous interviews, we heard that some found certain objects or technology to be helpful when trying to stop behaviors. For example, a toothbrush was offered to a girl who plucked her hair, and a favorite DVD was offered to a boy who liked to watch comedy. What do you use to manage [*or substitute other word they use*] this behavior?
        i. Are there objects you use to manage [*or substitute other word they use*] the behavior? (Further prompting: please describe any objects/technology you have seen used).
            1. Everyday/non-digital objects (ball, brush, blanket)?
            2. Digital technology (phones, computers)?
            3. Entertainment (video, DVD)?

    d. What are other approaches you are aware of?
        i. What about these worked well? What part is usually effective/smooth to implement?
            a. (*For further prompting*: Please describe examples of approaches that helped decrease self-injurious behavior).
        ii. What about these approaches could be better?
        iii. What may cause you to change your approach?
    e. What are examples of goals/hopeful outcomes for these approaches?

5. How do you track/log progress towards your goal/outcome?
    a. What makes it challenging to track/log these behaviors?

6. How do children respond to noise feedback, like an alarm?
    a. Visual feedback?
    b. Touch?
    c. Vibration?
    d. *For each sense, if unclear:* Are these responses positive (approach the sound/image etc.) or negative (retreat from the sound/image etc.)?

7. What effect does self-injurious behavior have on your classroom/educational setting?
    a. (*Prompt)* What challenges come from students who have self-injurious behavior?
    b. (*Prompt)* How do other students respond to students with self-injurious behavior?
    c. When do you work together with other educators to reduce self-injurious behavior?
    d. When do you work on your own?

8. When do you work together with parents to address SIB? When do you work separately/on your own? [*For further prompting on a and b, ask about challenges*]
    a. How is the home different from the school environment when managing SIB?
    b. What are challenges in transitioning school interventions to other environments (like the home)?

9. In what areas do you feel confident when managing self-injurious behavior?

10. What do you need help with when you try to manage self-injurious behavior?

11. How do you feel about using a technology that is not worn, and instead installed in an environment or a room to track SIB automatically?
    a. What would get in the way of applying something that is in the environment/room to track SIB?

12. How do you feel about asking children to wear something that would track SIB automatically?
    a. What would get in the way of applying something that could be worn to track SIB?
    b. [if need prompt] How would students react to wearing something added to his/her body that could automatically log behavior?
        i. Smart watch/bracelet?
        ii. Hat or headband?
        iii. Necklace?
        iv. Scarf?
        v. A sleeve/cuff?
        vi. Anklet?
        vii. Socks or shoes?
        viii. Belt?
        ix. Pants?
        x. Shirt?

xi. Any other articles of clothing or accessories that your child does or does not like?

13. [*Show technology, if in* person]
14. [Repeat questions 14, 15; add, "What do you like about this technology?" "What would be helpful about this technology?" "What do you not like?" "What would be challenging?" "When would [type of technology] be useful/not useful?"]

15. If a team of people were trying to create methods that could help self-injurious behavior in some way, what would you want them to know?
    a. What would you want them to know [about you, your educational environment, your students] if they are trying to create something wearable/that can be worn to track this behavior?
    b. What would you want them to know [about you, about your educational environment, about your students] if they are trying to create something that is not wearable – like something in the room/environment – to track this behavior?

Appendix B – Chapter 3

Table B1. Validation sample data mean classification performance across all participants, (0,1) labels.*

Algorithm	Mean Accuracy	Mean Specificity	Mean F-Score	Mean Train Time (kHz)
Linear SVM	0.820	0.770	0.829	5.15
Cubic SVM	0.905	0.882	0.903	0.329
Quadratic SVM	0.893	0.864	0.893	2.54
Fine Gaussian SVM	0.859	0.981	0.835	9.80
Med Gaussian SVM	0.907	0.889	0.903	13.6
Course Gaussian SVM	0.838	0.787	0.854	16.4
kNN, k = 1	0.934	0.924	0.936	53.8
kNN, k = 2	0.930	0.946	0.928	51.3
kNN, k = 3	0.917	0.883	0.922	59.2
kNN, k = 4	0.916	0.913	0.919	62.1
kNN, k = 5	0.913	0.881	0.914	65.4
kNN, k = 6	0.910	0.893	0.908	69.4
kNN, k = 7	0.905	0.869	0.905	59.5
kNN, k = 8	0.899	0.871	0.903	74.6
kNN, k = 9	0.901	0.858	0.905	71.9
kNN, k = 10	0.888	0.863	0.889	73.0
Weighted kNN	0.916	0.881	0.916	63.3
Cubic kNN	0.891	0.861	0.893	64.9
Cosine kNN	0.896	0.874	0.897	70.9
SRC	0.828	0.781	0.836	<u>.00182</u>
Fine DT	0.880	0.881	0.874	24.8
Medium DT	0.869	0.852	0.873	29.0
Course DT	0.819	0.780	0.825	43.1
Linear DA	0.802	0.756	0.809	59.5
Quadratic DA	0.845	0.792	0.850	44.8
Naïve Bayes	<u>0.754</u>	<u>0.643</u>	<u>0.778</u>	34.2
NN, layers = 1	0.829	0.786	0.838	1.28
NN, layers = 2	0.799	0.795	0.785	1.23
NN, layers = 5	0.848	0.799	0.853	1.47
NN, layers = 8	0.866	0.830	0.867	1.61
NN, layers = 10	0.882	0.859	0.880	1.74
NN, layers = 20	0.880	0.848	0.880	1.74
NN, layers = 30	0.878	0.857	0.879	1.77

*Note: <u>underlined</u> numbers indicates inferior performance measures compared to other methods.

Table B2. Mean kNN classification performance across all participants, for each behavior.*

Behavior	**Label**	**Acc**	**Spec**	**F**	**PredTime (kHz)**
Non-SIB events	0	0.915	0.961	0.874	5.56
Foot to surface	1	0.887	0.980	0.919	5.99
Fist to surface	2	0.894	0.978	0.909	6.62
Head hitting with hands/object	3	0.971	0.993	0.983	2.92
Picking fingers	4	0.953	0.967	0.964	4.85
Scratching	3	0.971	0.993	0.983	2.92
Throwing body on object/surface	5	0.953	0.994	0.984	2.40
Picking lip	6	0.885	0.966	0.884	32.3
Head banging (against wall)	7	0.970	0.992	0.985	19.2
Self-biting	8	0.952	0.992	0.957	15.9
Self-hitting	9	0.898	0.985	0.913	11.1
Pulling teeth (with hands)	10	0.898	0.991	0.887	11.1
Eye gauging	11	0.863	0.974	0.890	9.26
Jabbing pelvis	12	0.898	0.987	0.941	11.1
Jabbing throat	13	0.863	0.981	0.914	9.26
Knocking jaw	14	0.863	0.988	0.913	9.26
SMM - blowing on fingertips	15	0.863	0.988	0.913	9.26
Pulling teeth (with object)	16	0.731	0.936	0.774	17.1
SMM- spinning	17	0.731	0.968	0.787	17.1
SMM - flapping	18	0.731	0.966	0.653	17.1
SMM - jumping/flapping	19	0.833	0.975	0.834	15.2
SMM - jumping/spinning	20	0.731	0.984	0.607	17.1
Pulling ear	21	0.731	0.992	0.720	17.1
Pulling hair	22	0.934	0.985	0.881	13.7

Table B3. Confusion matrix for kNN predicted labels 0-23 group 10, all participants.*

Bx	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**
0	**28**	6	11	3	15	1	5	1	8	12	2	4	3	3	2	2	5	11	2	1	1	0	2	5
1	10	**69**	5	7	0	0	7	1	2	7	4	4	2	2	1	3	3	2	1	2	0	0	1	0
2	20	9	**42**	2	2	0	9	0	1	4	12	4	3	2	0	3	3	0	8	3	4	1	1	0
3	7	5	4	**88**	2	0	3	0	0	3	4	0	0	0	2	0	8	5	1	0	1	0	0	0
4	23	2	1	3	**90**	0	0	0	2	1	2	0	2	1	0	0	3	1	0	0	0	0	2	0
5	5	2	1	0	1	**120**	0	0	0	0	0	0	0	1	0	2	1	0	0	0	0	0	0	0
6	7	6	19	7	2	0	**56**	0	2	2	13	2	2	1	0	0	1	2	1	6	3	0	1	0
7	0	3	0	1	5	0	0	**122**	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
8	8	1	1	1	3	2	6	0	**92**	2	2	1	2	0	1	1	2	0	0	2	1	0	0	5
9	13	12	7	4	8	0	8	0	4	**49**	3	3	4	1	2	3	1	3	1	1	1	1	1	3
10	10	9	15	5	2	1	11	0	7	6	**36**	6	1	4	1	4	1	1	3	6	3	1	0	0
11	7	4	4	0	0	0	4	0	5	5	5	**89**	3	4	0	0	1	0	0	0	1	0	1	0
12	4	6	6	4	5	0	2	0	0	6	4	0	**85**	0	0	10	0	0	0	0	0	0	0	1
13	6	5	0	1	1	2	0	2	2	5	5	2	4	**90**	4	0	1	1	0	0	1	0	0	1
14	3	1	1	2	0	1	0	0	3	8	1	1	6	9	**88**	5	1	0	0	1	0	0	1	1
15	13	7	2	3	1	2	2	0	0	5	2	1	3	3	4	**82**	2	0	0	0	0	0	0	1
16	10	18	2	3	11	1	8	0	6	6	9	10	3	4	1	5	**29**	2	3	0	1	0	0	1
17	8	11	5	13	2	1	0	0	3	4	3	4	9	2	2	0	10	**52**	2	1	0	0	1	0
18	12	12	15	3	2	0	5	0	1	3	14	4	1	2	0	1	16	4	**30**	6	2	0	0	0
19	13	2	14	2	3	0	10	0	8	1	6	0	0	0	1	0	11	2	10	**35**	13	1	0	1
20	5	9	12	0	4	0	10	0	3	0	18	1	1	1	0	2	1	3	8	14	**36**	4	0	1
21	1	5	4	1	0	0	5	0	1	2	11	1	0	2	0	1	2	0	5	8	12	**11**	0	0
22	16	4	8	1	9	0	4	0	4	4	2	2	8	1	1	5	3	2	2	2	0	0	**54**	1
23	10	1	3	5	5	3	3	6	12	8	1	0	4	0	4	0	1	8	3	0	1	0	0	**55**

*Bx = behavior; row = actual and column = predicted

Appendix C – Chapter 4

Table C1. Means and standard deviations for select nonlinear variables. Variable names presented as "Variable[axis]". DFA = detrended fluctuation analysis exponent, saen = sample entropy, crossSaen = cross sample entropy, Rec = recurrence, Dtrmnsm = determinism, LdiagMAX = maximum diagonal length, MeanVertTT = trapping time, Div = divergence, Lmnrty = laminarity.

DFAx1	DFAy1	DFAz	saenX	saenY1	saenZ1	crossSaenXY1	crossSaenXZ1	crossSaenYZ1	RecX	DtrmnsmX
0.488	0.488	0.497	0.636	0.610	0.589	0.155	0.147	0.187	1.052	15.939
0.483	0.478	0.488	0.622	0.653	0.585	0.140	0.170	0.169	1.056	15.836
0.125	0.129	0.133	0.481	0.484	0.480	0.393	0.362	0.435	0.446	20.157
0.137	0.148	0.134	0.491	0.531	0.491	0.387	0.384	0.412	0.444	21.127

LdiagMAXX	MeanVertTTX	RecY	LdiagMAXY	DivY	MeanVertTTY	RecZ	DtrmnsmZ	LdiagMAXZ	DivZ	LmnrtyZ	MeanVertTTZ
2.942	1.316	1.068	2.927	0.117	1.345	1.080	17.653	3.328	0.119	5.926	1.390
3.050	1.269	1.038	2.751	0.100	1.173	1.072	17.626	3.177	0.115	5.425	1.271
4.036	1.597	0.453	4.201	0.133	1.587	0.440	21.372	4.557	0.131	9.770	1.609
4.620	1.580	0.457	4.232	0.129	1.554	0.441	22.085	4.405	0.131	9.494	1.576

Table C2. Loadings for the first 12 principal components (PC) extracted.

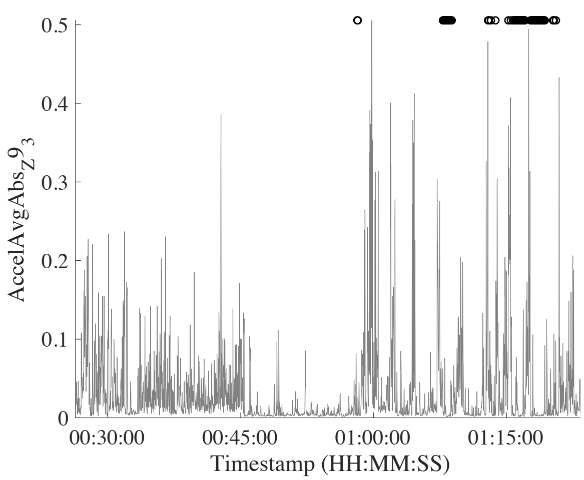
		Principal Components											
Type	**Feature**	**PC1**	**PC2**	**PC3**	**PC4**	**PC5**	**PC6**	**PC7**	**PC8**	**PC9**	**PC10**	**PC11**	**PC12**
Time Domain	**Correlation Coefficient XY**	0.023	0.012	-0.037	0.039	-0.131	-0.016	0.058	0.113	0.046	0.133	0.139	-0.232
	Correlation Coefficient XZ	0.013	0.003	-0.037	0.117	-0.107	0.040	-0.181	-0.051	-0.052	0.184	-0.051	-0.068
	Correlation Coefficient YZ	0.002	-0.076	0.041	-0.066	-0.045	-0.082	0.086	-0.004	-0.002	-0.022	-0.049	0.635
	Local Minima Count X	-0.135	0.137	0.139	0.122	-0.109	0.186	0.022	0.002	0.060	-0.030	-0.183	0.057
	Mean Absolute Value of X	0.040	-0.137	0.351	-0.193	-0.131	-0.031	0.109	-0.014	0.058	0.158	-0.256	-0.040
	Peak X	0.209	0.017	0.237	-0.061	-0.059	-0.016	0.061	0.050	0.032	0.057	-0.089	-0.037
	Minimum X	-0.117	-0.121	0.307	-0.167	-0.117	0.021	0.073	-0.013	0.073	0.176	-0.206	0.013
	10% X	-0.115	-0.139	0.216	-0.287	-0.098	0.018	-0.084	0.108	-0.016	-0.180	0.321	-0.041
	25% X	-0.071	-0.127	0.249	-0.302	-0.103	0.013	-0.082	0.121	-0.020	-0.197	0.344	-0.044
	99% X	0.161	-0.008	0.260	-0.199	-0.067	-0.009	-0.039	0.135	-0.014	-0.145	0.284	-0.033
	Jerk X	0.228	0.139	0.135	0.046	-0.010	0.040	0.006	0.049	0.007	0.019	0.005	0.009
	Variance X	0.181	0.159	0.144	0.020	0.009	0.040	-0.034	0.013	0.057	0.051	0.133	0.170

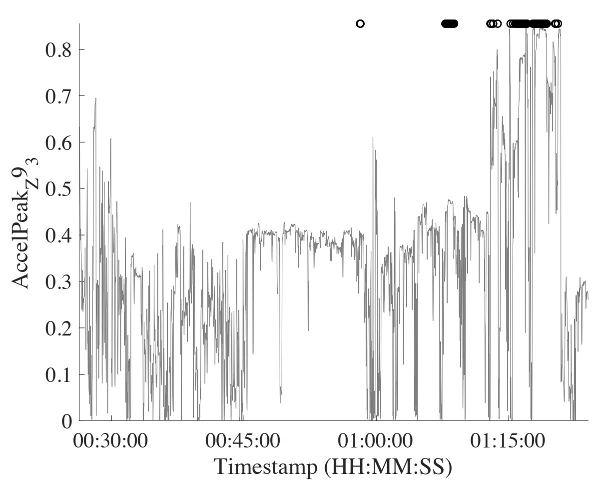
	Local Minima Count Y	-0.135	0.143	0.153	0.097	-0.020	0.165	0.059	0.112	0.045	-0.076	-0.118	0.122
	Zero Crossings Y	0.121	0.113	0.048	-0.119	-0.124	-0.014	0.102	-0.039	-0.064	0.076	-0.227	-0.160
	Mean Absolute Value of Y	0.054	-0.054	0.015	0.414	0.192	0.015	-0.113	0.061	0.029	-0.111	0.314	0.143
	1% Y	-0.147	-0.004	0.061	-0.186	0.355	0.083	-0.280	-0.102	-0.112	0.038	-0.102	-0.019
	50% Y	-0.009	0.063	0.118	-0.186	0.432	0.096	-0.343	-0.147	-0.132	0.021	-0.130	-0.042
	99% Y	0.153	0.119	0.145	-0.103	0.321	0.081	-0.274	-0.117	-0.078	-0.002	-0.063	-0.001
	Jerk Y	0.222	0.149	0.118	0.093	-0.001	0.058	-0.005	0.005	0.018	-0.020	0.010	0.002
	Variance Z	0.184	0.170	0.149	0.024	-0.001	0.072	0.023	0.019	0.050	0.044	0.074	0.159
	Zero Crossings Z	0.138	-0.051	0.147	0.208	0.110	-0.071	0.027	0.000	0.031	0.064	-0.033	-0.049
	Mean Absolute Value of Z	0.029	0.314	-0.249	-0.254	-0.116	0.061	0.014	-0.011	-0.040	-0.009	0.051	0.026
	RMS Z	0.067	0.322	-0.223	-0.236	-0.110	0.064	0.017	-0.012	-0.036	-0.012	0.051	0.036
	Peak Z	0.200	0.238	-0.031	-0.073	-0.072	0.073	-0.002	-0.008	-0.011	-0.038	0.020	0.037
	Minimum Z	-0.126	0.249	-0.212	-0.225	-0.056	0.065	0.009	0.029	-0.005	0.061	0.090	0.040
	50% Z	-0.042	0.052	-0.037	-0.170	0.392	-0.094	0.395	0.105	0.260	-0.060	-0.018	-0.042
	90% Z	0.032	0.056	-0.017	-0.154	0.392	-0.099	0.415	0.087	0.254	-0.069	-0.029	-0.038
	Jerk Z	0.218	0.167	0.116	0.068	-0.036	0.063	-0.004	0.010	0.011	-0.017	-0.018	-0.004
Frequency Domain	**First FFT Amplitude X**	0.216	0.106	0.056	0.042	0.035	-0.039	0.043	0.019	-0.010	0.030	0.030	0.011
	Second FFT Amplitude X	0.234	0.110	0.095	0.029	0.024	-0.015	0.019	0.041	0.009	0.051	0.069	0.018
	Second FFT Peak X	-0.049	0.032	0.111	0.049	-0.100	0.091	-0.026	0.080	0.013	-0.136	-0.221	-0.022
	First FFT Amplitude Y	0.211	0.076	0.044	0.000	-0.011	-0.053	-0.030	-0.065	0.011	0.044	0.034	0.032
	First FFT Peak Y	-0.098	0.079	0.078	0.189	0.076	0.181	-0.023	0.153	0.101	-0.186	0.046	-0.124
	Second FFT Peak Y	-0.054	0.061	0.059	0.168	-0.007	0.179	-0.023	0.134	0.057	-0.261	-0.078	-0.118
	First FFT Peak Z	-0.078	0.188	-0.036	-0.011	-0.007	0.184	-0.075	0.148	0.101	-0.130	-0.052	-0.186
	Second FFT Peak Z	-0.045	0.125	0.011	0.023	-0.052	0.131	-0.018	0.131	0.040	-0.293	-0.171	-0.164
Nonlinear Variability	**DFA Alpha X**	0.183	-0.041	-0.107	-0.014	0.003	-0.126	0.025	-0.049	-0.096	-0.156	-0.051	-0.041
	DFA Alpha Y	0.195	-0.045	-0.080	-0.032	-0.068	-0.098	-0.026	-0.115	-0.048	-0.105	-0.069	-0.053
	DFA Alpha Z	0.175	-0.128	-0.009	0.026	-0.018	-0.095	0.046	-0.116	-0.069	-0.170	-0.112	-0.019
	Sample Entropy X	-0.096	0.079	0.095	0.056	0.020	0.037	0.022	-0.017	0.104	0.369	0.123	-0.105
	Sample Entropy Y	-0.098	0.079	0.100	0.091	0.070	0.044	0.058	0.041	0.044	0.317	0.124	-0.020
	Sample Entropy Z	-0.089	0.167	0.019	0.036	0.027	0.018	-0.023	0.055	0.061	0.369	0.178	-0.084
	Cross Sample Entropy XY	0.111	0.068	-0.067	-0.032	-0.011	-0.075	-0.049	-0.100	-0.058	-0.027	-0.085	-0.011
	Cross Sample Entropy XZ	0.112	-0.031	-0.009	0.079	0.125	-0.085	0.106	-0.005	0.012	-0.107	0.042	-0.064
	Cross Sample Entropy YZ	0.099	-0.043	0.047	0.000	-0.072	-0.026	0.049	-0.093	-0.001	-0.020	-0.084	-0.326
	Recurrence X	0.049	-0.069	-0.077	-0.030	0.011	0.014	-0.112	0.495	-0.071	0.147	-0.189	0.130
	Determinism X	0.160	-0.132	-0.136	-0.054	0.068	-0.037	-0.095	0.333	-0.049	0.096	0.009	-0.071
	Maximum Diagonal Length X	0.147	-0.125	-0.123	-0.052	0.060	-0.026	-0.100	0.386	-0.069	0.119	-0.022	-0.057
	Trapping Time X	0.088	-0.098	-0.094	-0.060	0.034	0.043	-0.082	0.374	-0.044	0.024	-0.202	0.156
	Recurrence Y	0.044	-0.067	-0.053	-0.035	-0.056	0.061	-0.233	-0.051	0.581	0.000	-0.044	0.075
	Maximum Diagonal Length Y	0.146	-0.113	-0.101	-0.063	-0.045	-0.017	-0.166	-0.031	0.293	0.058	0.023	-0.065
	Divergence Y	0.083	-0.098	-0.120	-0.066	-0.043	-0.034	-0.151	-0.094	0.257	0.010	0.000	-0.091

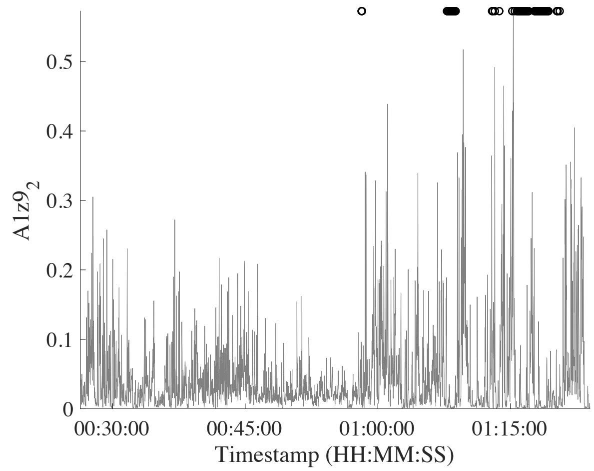
	Trapping Time Y	0.090	-0.098	-0.095	-0.054	-0.059	0.058	-0.190	-0.084	0.461	-0.035	-0.043	0.079
	Recurrence Z	0.041	-0.120	-0.043	-0.027	0.022	0.429	0.140	-0.072	-0.031	0.078	0.048	-0.045
	Determinism Z	0.157	-0.213	-0.080	-0.013	0.034	0.191	0.084	-0.013	-0.062	0.053	0.055	-0.138
	Maximum Diagonal Length Z	0.147	-0.185	-0.079	-0.010	0.031	0.229	0.091	-0.004	-0.065	0.078	0.050	-0.189
	Divergence Z	0.079	-0.154	-0.036	-0.034	0.023	0.116	0.092	-0.151	-0.085	-0.024	0.071	0.053
	Laminarity Z	0.099	-0.129	-0.066	-0.046	0.014	0.421	0.119	-0.041	-0.011	0.013	-0.004	0.143
	Trapping Time Z	0.081	-0.151	-0.073	-0.050	0.013	0.436	0.138	-0.106	-0.044	0.032	0.036	0.142
Explained Variance (%)		**23.011**	**8.014**	**6.323**	**5.270**	**3.900**	**3.562**	**3.249**	**3.006**	**2.624**	**2.455**	**2.269**	**1.876**

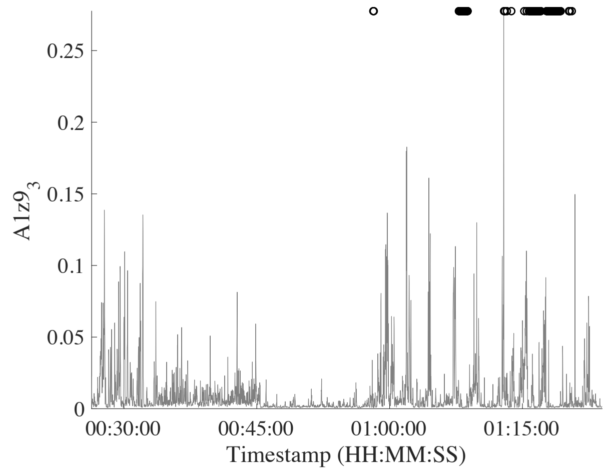
Sample participant feature data

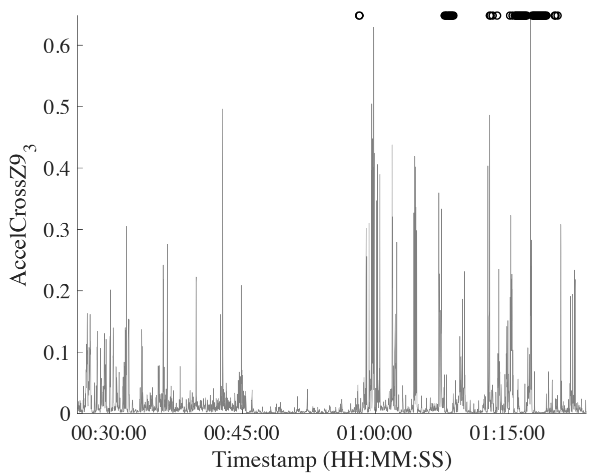
Features with labeled SIB (black circles at the top of the graph). Feature$_c$P$_S$ = Feature from c = channel (XYZ), P = participant number, and s = sensor (2 = wrist, 3 = waist). Accel = Acceleration (in G); AvgAbs = mean absolute value; A1 = amplitude of first FFT peak; Cross = zero crossings; RecRate = recurrence rate; Dtrmnsm = determinism; MeanVertTT = trapping time; crossSaen = cross-sample entropy

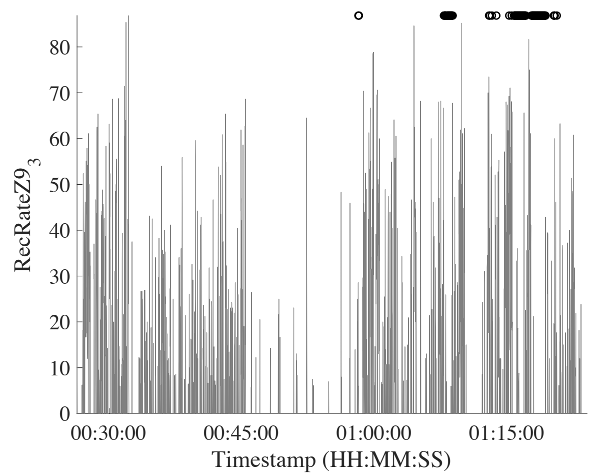


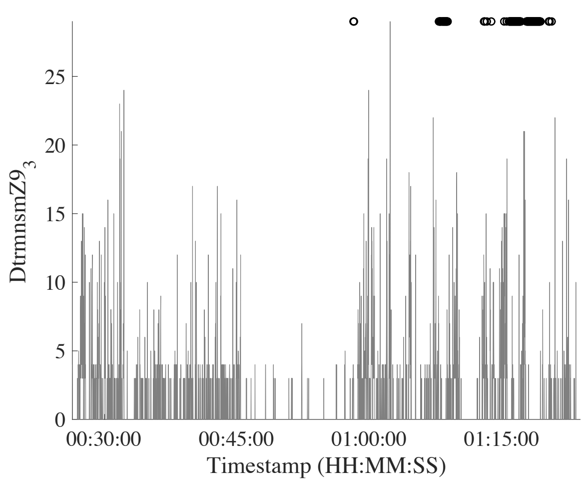


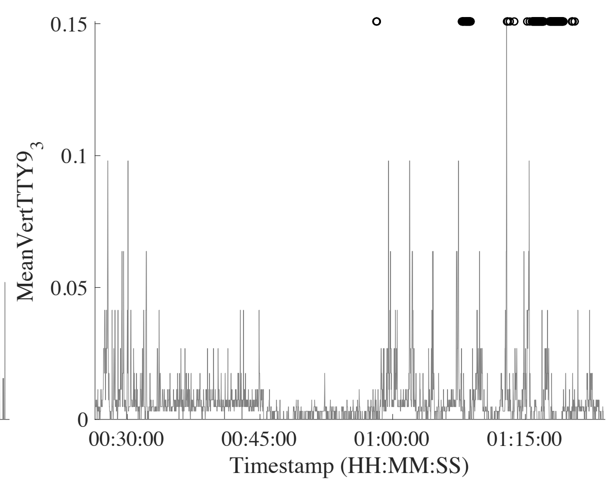

A1z9$_2$
0
0.1
0.2
0.3
0.4
0.5
00:30:00
00:45:00
01:00:00
01:15:00
Timestamp (HH:MM:SS)


A1z9$_3$
0
0.05
0.1
0.15
0.2
0.25
00:30:00
00:45:00
01:00:00
01:15:00
Timestamp (HH:MM:SS)


AccelCrossZ9$_3$
0
0.1
0.2
0.3
0.4
0.5
0.6
00:30:00
00:45:00
01:00:00
01:15:00
Timestamp (HH:MM:SS)


RecRateZ9$_3$
0
10
20
30
40
50
60
70
80
00:30:00
00:45:00
01:00:00
01:15:00
Timestamp (HH:MM:SS)


DtrmnsmZ9$_3$
0
5
10
15
20
25
00:30:00
00:45:00
01:00:00
01:15:00
Timestamp (HH:MM:SS)


MeanVertTTY9$_3$
0
0.05
0.1
0.15
00:30:00
00:45:00
01:00:00
01:15:00
Timestamp (HH:MM:SS)

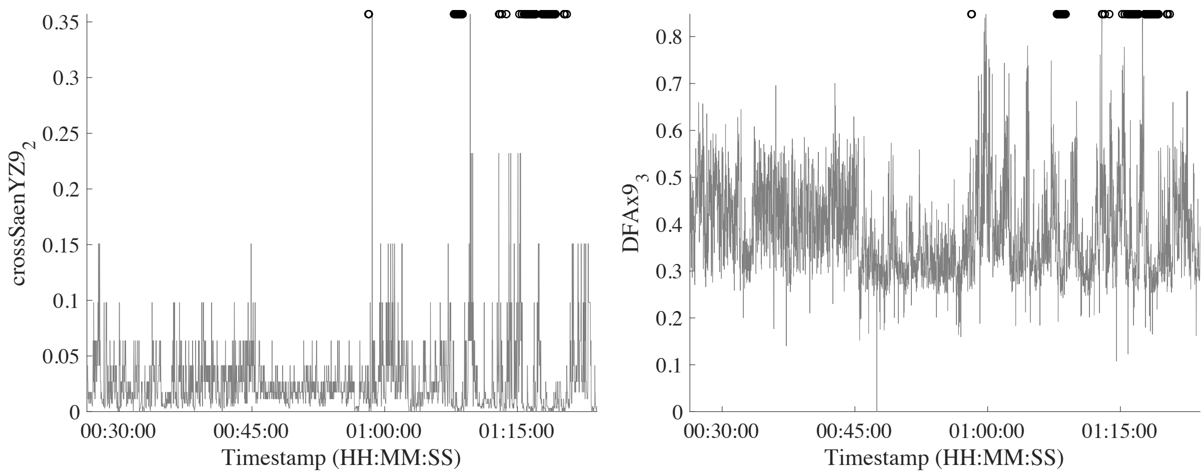


Table C3. Descriptive statistics for head banging versus non-SIB events for P9. DFA = detrended fluctuation analysis exponent, saen = sample entropy, crossSaen = cross sample entropy, RecRate = recurrence rate (%), Dtrmnsm = determinism (%), LdiagMAX = maximum diagonal length, Div = divergence (%), Lmnrty = laminarity (%), MeanVertTTX = trapping time. Variables are formatted as "VariableName[channel][participant number]_[sensor number]". Sensor number 3 = waist.

	DFAx9_3	DFAy9_3	DFAz9_3	saenX9_3	saenY9_3	saenZ9_3	crossSaenXZ9_3	crossSaenYZ9_3
Mean (Non-SIB)	0.656	0.372	0.350	0.393	0.843	0.891	0.006	0.060
Mean (SIB)	0.457	0.441	0.482	0.567	0.612	0.581	0.038	0.203
SD (Non-SIB)	1.332	0.085	0.080	0.101	0.514	0.505	0.078	0.254
SD (SIB)	0.127	0.072	0.116	0.467	0.459	0.469	0.156	0.399
Min (Non-SIB)	0.070	0.000	0.000	0.000	-0.062	-0.061	0.000	-0.072
Min (SIB)	0.280	0.308	0.322	0.000	0.000	0.150	0.000	0.000
Max (Non-SIB)	21.836	0.805	0.872	0.848	3.080	3.235	1.774	2.793
Max (SIB)	0.805	0.527	0.728	1.483	1.512	1.848	0.643	1.390

	RecRateX9_3	DtrmnsmX9_3	LdiagMAXX9_3	DivX9_3	LmnrtyX9_3	MeanVertTTX9_3
Mean (Non-SIB)	2.234	1.049	6.890	1.275	0.121	3.066
Mean (SIB)	1.011	3.587	0.765	0.074	1.740	0.706
SD (Non-SIB)	2.542	0.463	14.438	2.770	0.356	6.862
SD (SIB)	0.564	7.805	1.437	0.138	3.255	1.312
Min (Non-	1.000	0.000	0.000	0.000	0.000	0.000

SIB)						
Min (SIB)	0.256	0.000	0.000	0.000	0.000	0.000
Max (Non-SIB)	42.771	1.905	77.273	28.000	2.322	60.976
Max (SIB)	1.868	27.500	4.000	0.333	8.333	3.000

	RecRateY9_3	DtrmnsmY9_3	LdiagMAXY9_3	DivY9_3	LmnrtyY9_3	MeanVertTTY9_3
Mean (Non-SIB)	4.459	1.039	3.748	0.744	0.062	2.073
Mean (SIB)	0.995	9.822	2.118	0.056	2.435	1.078
SD (Non-SIB)	7.146	0.470	10.455	2.098	0.258	5.539
SD (SIB)	0.596	18.434	4.256	0.104	3.855	1.507
Min (Non-SIB)	1.019	0.037	0.000	0.000	0.000	0.000
Min (SIB)	0.220	0.000	0.000	0.000	0.000	0.000
Max (Non-SIB)	70.714	1.905	76.471	25.000	2.000	45.455
Max (SIB)	1.868	59.459	16.000	0.333	11.905	3.333

	RecRateZ9_3	DtrmnsmZ9_3	LdiagMAXZ9_3	DivZ9_3	LmnrtyZ9_3	MeanVertTTZ9_3
Mean (Non-SIB)	2.030	1.049	7.351	1.376	0.131	3.215
Mean (SIB)	0.907	16.283	3.118	0.191	1.875	0.424
SD (Non-SIB)	1.516	0.463	15.445	3.054	0.374	7.183
SD (SIB)	0.387	16.019	2.913	0.147	6.339	1.204
Min (Non-SIB)	1.000	0.000	0.000	0.000	0.000	0.000
Min (SIB)	0.256	0.000	0.000	0.000	0.000	0.000
Max (Non-SIB)	17.133	1.905	86.842	29.000	2.000	62.903
Max (SIB)	1.575	60.606	10.000	0.333	25.806	4.000