Evaluation of the Effect of Approximate Utterance Reward Stimulation in a Game for ASD

James Lyons-Weiler, PhD

Institute for Pure and Applied Knowledge

DISCLOSURE: This proposal is being written at the request of Nea Hanscomb. However, Ms. Hanscomb and her affiliates have no influence over the direction or outcomes of the research, and Dr. Lyons-Weiler has no vested financial interest in the product being evaluated. All results will be published, positive or negative, and feedback provided by IPAK to Ms. Hanscomb and affliates on application design features are given to Ms. Hanscomb free of charge by Dr. Lyons-Weiler and IPAK consultants. No pro-bono relationship whatsoever exists, and no long-term financial reward related to profits from the sale or licensing of Say It! SAM has been offered to IPAK, Dr. Lyons-Weiler, and IPAK Consultants from anyone, including Say It! SAM, Technology Access, INC, or WORDS+ or any of their affiliates. The success or failure of the proposed research is defined as "completion of the study and testing for a benefit of AURS". The effort to produce this proposal was funded by a personal donation to IPAK from Ms. Hanscomb.

BACKGROUND AND SIGNIFICANCE

The rates of autism have increased from 1 in 100,000 in the 1970's to 1 in 68 in 2008 (CDC). Typically, autism involves developmental delays three of five domains (Table 1). An entire industry of Applied Behavioral Analysis therapy (ABA) has accompanied this increase. ABA intensity (frequency of sessions per week) has been found to be secondary to the duration (number of total session) as a predictor of improvement in language skills. The detected improvements have involved a large number of diverse measures, which provide a rich source of outcome variables. This is especially true given the advancement in ABA is the Pivot Reward System (National Professional Development Center on Autism Spectrum Disorders, 2010), in which the activities are determined by client preferences. It is believed that pivoting provides a sense of reward providing a sense of control, allow the clients to develop an internal sense of control over their own developing language skills. ABA requires many hours per client per week, and any technology that could be found to increase the rate of improvement in language skills in autistics would have many benefits to society.

Apraxia occurs in a large percentage of children with autism. A keystone study on predictors of speech development conducted by Wodka et al. (2013) found that age, internalizing symptoms, higher non-verbal IQ and lower social impairment to be predictors of phrase speech development, and that age, higher non-verbal IQ and lower social impairment were good predictors of fluent speech development in children with ASD. Importantly, half of the non-speaking children in the study eventually development speech.

An observational study of spontaneous language samples led Wittke et al. (2017) to identify three subgroups based on language ability: subgroups: those with normal language, those with marked difficulty in grammatical production but relatively intact vocabulary, and those with more globally low language abilities. Mottron (2017) argues that a primary focus on developing language skills should be a priority in interventions in ASD due to likely consequential improvements in other developmental domains, and that the role of passive, intrinsic learning has been underestimated. He stressed that the

combination of visual and auditory stimuli may better match the modes of language learning in ASD, and the non-captive settings and a sense of autonomy may be expected to accelerate otherwise delayed

Core Symptoms	Key Symptoms	Frequently Observed Conditions
Language delay	Stereotyped behaviors, Hyperactivity	Inattention, obsessions
Lack of social/ emotional reciprocity	Insistence on sameness	Compulsions, picky eating
Lack of nonverbal communication	Insistence on sameness	Intellectual disabilities
Lack of developing/ maintaining	Sensory variance	Relationships

Table 1. Core and key symptoms and frequently observed effected conditions in ASD (Source: Lyons-Weiler, 2016)

language acquisition. Age of first words is good predictor of language development outcomes in non-regressive cases of autism (Mayo et al., 2013). Assessment of both non-regressive and regressive toddlers with ASD by Kover et al. (2016) found that components of linear models of *age of first words, age of first phrase*, and *number of words* both strongly predicted language development progress (receptive and expressive), non-verbal cognition, and, to lesser degree, adaptive behavior. In another study (Ellis Weismer and Kover, 2015), the same team found that *cognition*, *maternal education*, and *response to joint attention* correctly classified 80% children who developed high language ability compared to low language ability. No independent test set has been used to evaluate generalizability of past model predictions.

Video games utilizing speech recognition have recently demonstrated a positive effect on speech production in speech therapy settings. Research conducted by Rubin et al. (2016, 2017) found that introducing speech-powered mobile video games into cleft speech therapy increased effort, therapy compliance, and therapist feedback. The same games were found to promote short-term speech production in individuals with developmental disabilities such as cerebral palsy. The team also found existing speech recognition could detect specific impairments and approximations with 95% accuracy for dictionaries under 100 words. Dr. Rubin is a co-Principle Investigator on the currently proposed study.

A recently developed software application called *Sayin' it Sam™* by Appropos Software is based on the principal of sensory reward for approximate utterances. In this setting, the app runs on a tablet and is presented to the client as a simple video game. Presented with an object, or a color, the client is rewarded with a preferred tone if the app detects and utterance approximating the correct English word. For example, if presented with a ball, and the client issues a verbal attempt beginning with "b", or a sound approximating "b", the app provides the sensory reward. For this proposal, we call this approach AURS (Approximate Utterance Reward Stimulation). Importantly, the stimulus for speech is external: the software prompts the user to speak the appropriate term associated with the image,

capitalizing on multi-sensory input (auditory stimulation and prompt), visual presentation of the object, autonomous control and sensory reward.

We are interested in exploring, via rigorous study, whether AURS augments ABA by enhancing language pathways via sensory rewards. To this end, we propose to conduct the following **Specific Aims:**

Specific Aim 1. We will conduct a retrospective Observational Study to establish a baseline of performance with, and without Approximate Utterance Reward Stimulation (AURS). At least 60 Applied Behavioral Therapists will be solicited to incorporate the AURS protocol into the ABA for two 4-year old clients each (one male, one female). The goal of this observational study will be to establish preliminary data characteristic of language development progression associated with the use of AURS in the ABA setting. Each ASD Research Outcomes Certified ABA Therapist will also be solicited to choose, at random, four "control" clients with ASD without intellectual disability (aged 4) and asked to report the same language progression outcome measures. Baseline data on non-verbal IQ, social impairment, and internalizing symptoms will be collected for each study participant. Baseline approximate utterance rates will be collected automatically from the software, per client. The study will run for 8 weeks. Data will be via the web aggregated and analyzed for a variety of measures typically available from progress reports. Primary and secondary analyses will include Self-Controlled Case Series analysis, ANOVA (one-way and two-way) and ANCOVA as appropriate. False discovery and family-wise error rate control will be applied per multiple comparisons as appropriate (Holm's Method and Benjamini-Hochberg method, respectively). This aim will allow us to (1) estimate the effect sizes for various outcome measures stratified by gender, (2) determine important variables that impact the utility of AURS, (3) and will allow us to determine an appropriate sample size for the randomized prospective study described as Specific Aim 2. Stratified analysis based on non-verbal IQ and social impairment scores will be conducted, and before-and-after scores for non-verbal IQ and social impairment will also be compared. Overall ADOS severity at the time of enrollment will also be considered via stratified (high/low) or via covariate analysis in ANCOVA linear modeling.

Estimated Total Cost: $100,000

Specific Aim 2. Randomized Prospective Study. In this study, *N* ASD clients without ID (aged 3-5) will be randomized to each arm of a study in which AURS is used by *M* ABA Therapists, with N and M determined via power analyses provided by Study 1 (Specific Aim 1). The same baseline data (non-verbal IQ, internalizing symptoms scores, social impairment, and approximate utterance rates) resulting from Specific Aim 1 and Specific Aim 2 will be combined to allow an intermediate Self-Controlled Case Series study that will inform us of the utility mid-study (8 weeks). AURS will be applied for a total of 16 weeks. At the end of 16 weeks, all outcome measures will be analyzed, stratified by gender using ANOVA (one-way and two-way, as appropriate) and ANCOVA as appropriate per primary and secondary analyses, as described in Specific Aim 1. Multiple comparison concerns will be addressed as in Specific Aim: Family-wise error rates will be controlled using Holm's Method; False-Discovery rates will be controlled using the Benjamini-Hochberg method. An ABA Therapist survey (Appendix 1) will be conducted to provide feedback from the clinician end-user's perspective.

Estimated cost: $150,000

Specific Aim 3. Inclusion criteria in Specific Aims 1 and 2 were restricted to control variability associated with age. Specific Aim 3 will allow us to study the effects of age of clients, and help us identify optimal

age-specific details of the AURS paradigm, and obtain further feedback on updates, options and upgrades. Features solicited from by the ABA Therapist population will have been incorporated and explored and the AURS approach will updated accordingly. A larger randomized prospective study will be conducted using the augmented software with age-appropriate content and sensory rewards. One-thousand ASD patients will be studied comparing the original AURS software interface and algorithm and AURS 2.0, augmented with content desirable to different ages. Patients aged 3-17 will randomized to two arms (AURS 2.0 vs. AURS 1.0) to determine the effect of the use of AURS in combination with ABA. Importantly, in contrast to Specific Aims 1 and 2, inclusion criteria will now allow intellectual disability, which will lead to a subset analysis and help seed further studies aimed specifically at clients with ASD with ID. Longitudinal responses of both AURS 1.0 and AURS will be studied over the entire age range to assess age-specific efficacy and responsiveness to alternative sensory rewards.

Estimated Total Cost: $250,000

STUDY DESIGN (Specific Aim 1)

Specific Aim 1 will be conducted as a retrospective study on prospectively collected data. Data collected from both arms (AURS vs. no AURS) will allow us to identify unanticipated source of random and systematic variation, including identifying potential confounders and determine appropriate samples sizes for sufficient statistical power to detect differences in outcome variables.

DATA ANALYSIS PLAN (Specific Aim 1)

In this phase, the analyses are, by design, exploratory. With systematic data reporting from ABA Therapists, we will examine retrospectively trends over time in search of evidence of satiation and decay effects, diminishing returns on outcomes, and overall temporal variability within and between groups. The primary analysis will be a Self-Controlled Case Series Study (SCCS; Whitaker et al., 2005), which compares final outcome measures to baseline, using each client as their own control. SCCS design provides estimates of the relative incidence, and effect sizes suitable for use for power analysis for follow-up studies. It will allow a preliminary test the main effect of exposure to AURS during weeks 2-7. The design has both strengths and weaknesses. Our primary concern is incidental temporal confounding variation unique to one arm. The full final analysis will include cross-group comparisons consistent with prospective randomized clinical trial standards as the Primary Analysis to avoid temporal confounding.

Primary Analysis: ANOVA and repeated measures ANOVA will be performed on outcome variables collected for all patients. False discovery and Family-wise error rates will be controlled as described in the Aims. Secondary Analyses: Two-way ANOVA will be used to study the effects of gender of both the clients (which will be balanced by design). A variety of secondary analyses will look for interactions between variables to test for effects of ABA Therapist age and experience.

Outcomes

The training module for ASD Research Competency be developed will include many nominal tracking variables directly reflecting speech so the competency includes a priori introduction to the specific definitions of the variables. These will include the Children's Communication Checklist (CCC2-NL, in English) and Vineland Adaptive Behavior Scales, Third Edition (Vineland-3) devices, which, combined, collect many independent variables within each domain for use as outcomes individually, or as composite scores for each domain assessed (Communication, Daily Living Skills, Motor Skills,

Maladaptive Behavior). The use of both instruments will allow the full spectrum of language ability within ASD to be assessed from non-verbal to speaking with sentences.

Ethical Considerations

To insure full transparency of trial reporting, we will utilize the industry-standard formatted figures reflecting accrual rates, and fates of clients as the study progresses. Our Study Report will include a full accounting of the fates of all patients allocated to both ABT+AURS and ABT-only (Control) groups (Fig 1).

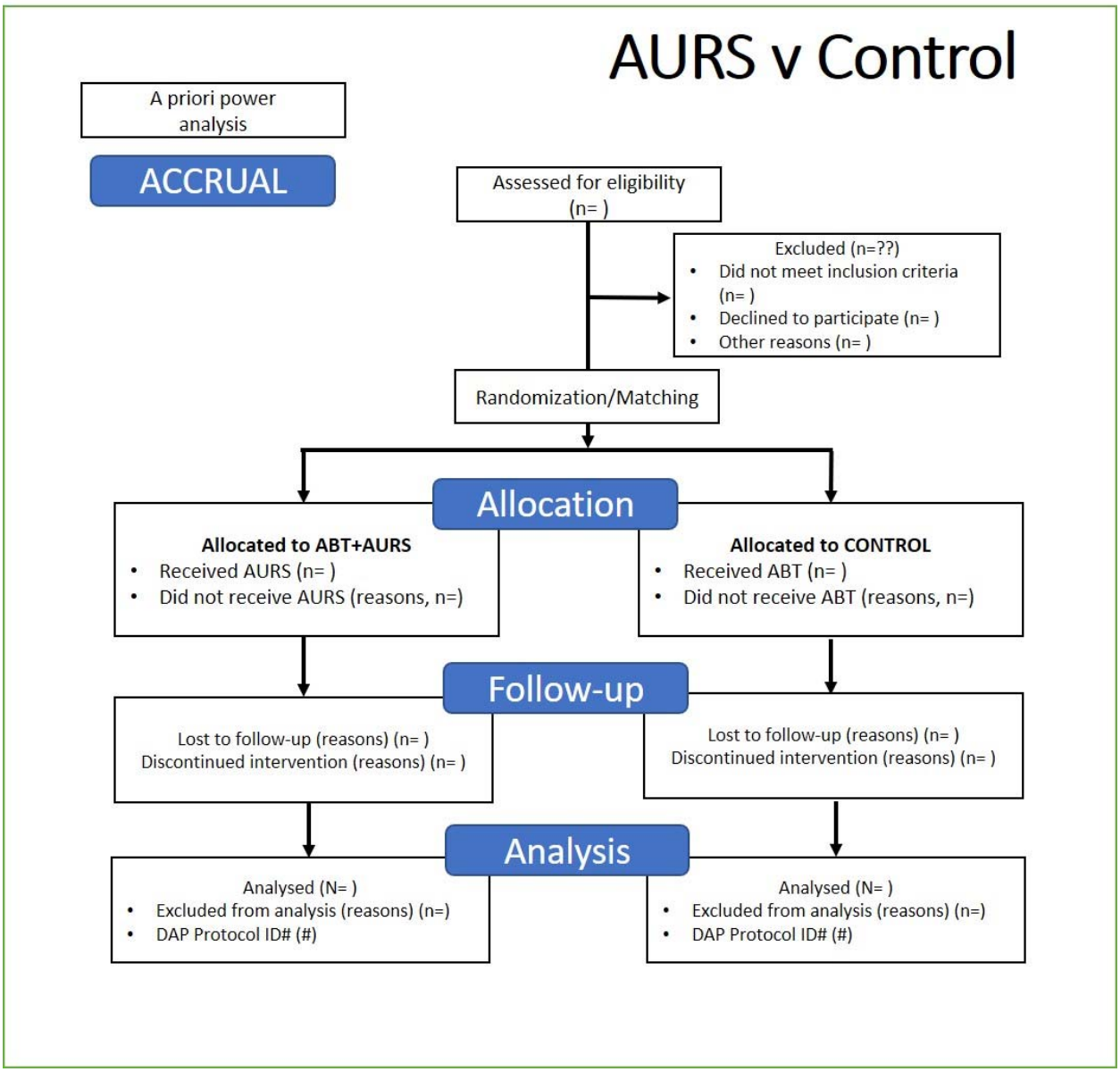


Figure 1. Patient Fates Tracking Form that will be used to insure transparency and objectivity of the results of our proposed study.

Potential Pitfalls

The benefits of using ABA Therapists working in the field as opposed to bringing clients into one center is that the results of our study, positive, negative, or null, will likely generalize the real-world clinical

setting. The clients involve will represent a randomized sample of a broad and diverse clinical population. The age restriction will control for variability associated with client age.

All ABA Therapists will be given three sets of instructions designed to minimize random sources of error and bias. They will be instructed how to select matched candidates for overall inclusion in the study in a randomized manner (Appendix 2), how to obtain informed consent for each client from parents

(Appendix 3), and training on outcomes evaluation (Appendix 4). They will acknowledge that their compensation (Amazon Gift Cards) are for the data reporting effort, not for "proving" efficacy of AURS (Appendix 5), and that the data will be checked for entry errors and bias. To avoid selection biases, assignment of consented clients to AURS or CONTROL group will be conducted by the investigators, not by the ABA Therapists. Age and gender of the ABA Therapists will be recorded, as will number of months performing ABA Therapy to explore any effect of ABA Therapist on the data. ABA Therapists may be novices to the assessment of non-verbal IQ, internalizing symptoms and social impairment. This will be addressed by proficiency testing during the ABA Therapist training period.

We are aware that the effective sample size (N) for the study design is the number of ABA Therapists, not the total number of clients. The degrees of freedom in ANOVA will executed in a manner that accommodates the effects of both sample sizes.

STUDY DESIGN (SPECIFIC AIM 2)

In contrast to the Retrospective Observational Study in Specific Aim 1, the study in Specific Aim 2 is to conducted as a Randomized Prospective Study, predicated and adjusted as necessary for design and implementation based on the results learning from the prior study. We expect to learn the sample size necessary for sufficient power to detect an effect of AURS, the most responsive outcomes variables to focus upon, and variables for which matching might be considered across treatment and control groups to reduce any effect of incidental confounding. For example, if a gender of ABA Therapist effect is found, we will want to be sure to match across AURS and CONTROL groups, or plan to stratify for ABA Therapist gender in the final Data Analysis Plan.

Data Analysis Plan (Specific Aim 2)

Unlike the pilot study in Specific Aim 1, no data will be analyzed in real time during the Prospective Study proposed in this aim. Instead, a Preliminary Report will be conducted using the Self-Controlled Case Series (SCCS) study design and analysis paradigm; see review by (Petersen et al. (2016). While SCCS has high intrinsic power, because each clients' baseline is used as its own control, it does not protect against incidental confounding that could occur due to unanticipated sources of bias. Our matching and randomization study design should protect against such problems; however, the study will continue to completion at 16 weeks. This will allow us to assess duration effects. In the final analysis, ANOVA, ANCOVA and multiple linear regression will be used, as appropriate, upon the 2, 4, 8, 12 and 16-week data. False discovery and family-wise error rates will be controlled as described in Specific Aim 1. Longitudinal data will be studied after the completion of the Prospective Study. The main effect of AURS and its interactions with gender of client will be assessed using the standard F-statistical. Adjustments for covariates learned in Specific Aim 1 that cannot be balanced via matching during client accrual may represent confounders, and will either be addressed statistically, or via stratified subgroup analysis.

Outcomes Measures and Social Validity Assessment (Aims 1-3)

We will closely follow the same outcomes measured in a study used to assess Pivotal Response Treatment by Verschuur et al. (2017). Our adoption of these measures will ensure comparability across studies.

Changes in the children's language skills, pragmatic skills, adaptive skills, and maladaptive behaviors, additional measures will be collected during baseline and post-intervention. The CCC2 will be used to

measure language skills and pragmatic skills. The CCC2-NL is a 70-item questionnaire designed to measure both structural and pragmatic aspects of children's language skills (Bishop 2003; Geurts 2007). The CCC2-NL consists of ten subscales: (a) speech, (b) syntax, (c) semantics, (d) coherence, (e) inappropriate initiation, (f) stereotyped language, (g) use of context, (h) nonverbal communication, (i) social relations, and (j) interests. Based on the subscales, three summary measures can be calculated: (1) general communication composite, indicating the child's communicative competence, (2) a social-interaction deviance composite, indicating the extent of social communication difficulties versus structural language deficits, and (3) a pragmatic composite, indicating the child's pragmatic abilities. High scores of the summary measures reflect higher impairment. In the present study, the general communication composite will be used to measure language skills; the pragmatic composite will be used to measure pragmatic skills. The CCC2-NL has demonstrated convergent validity, internal consistency, and test–retest reliability, and is sensitive enough to distinguish between children with ASD, specific language impairments and attention-deficit/ hyperactivity disorder (Geurts 2007).

Source: http://www.pearsonclinical.com/language/products/100000193/childrens-communication-checklist2-us-edition-ccc-2.html

The Vineland 3 is a standardized assessment of adaptive behavior and provides standard scores on four domains: communication, daily living skills, socialization, and motor skills (Sparrow et al. 2005). Furthermore, the Vineland-3 provides an overall standard score: the adaptive behavior composite (ABC). The Vineland-3 also provides a maladaptive behavior index (MBI), a composite of internalizing, externalizing and other maladaptive behaviors that may interfere with the individual's adaptive functioning. In the present study, the ABC and the standard scores on communication, daily living skills and, socialization will be used to measure adaptive skills. Higher scores indicate higher levels of adaptive functioning; lower scores indicate lower levels of adaptive functioning. The MBI will be used to measure maladaptive behaviors. Higher scores indicate higher levels of maladaptive behavior; lower scores indicate lower levels of maladaptive behavior. During each week of intervention, ABA Therapists members will be asked to fill in the Vineland 3 parent/caregiver rating form for their clients.

Source: https://images.pearsonclinical.com/images/Assets/vineland-3/Vineland-3-Flyer.pdf

Social Validity

During post-intervention, credentialed staff members will be asked to fill in a questionnaire to assess the social validity of AURS in general, and of the AURS use training prior to the initiation of the study. The questionnaire will consist of various statements (e.g., 'I am willing to use AURS during ABA of my clients' and 'The instructions on application of AURS informative') that will be rated on a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The questionnaire will be used to measure ABA Therapists' attitude towards AURA and whether they considered the components of the AURS staff training as effective, relevant and pleasant.

Data Sharing Plan

All de-identified data from each Specific Aim will be published at the time of the publication of the study report(s) (i.e., peer-reviewed publications) resulting from that aim.

Future Studies

Early intervention is known to be a factor in the efficacy of ABA. Studies designed to learn age-appropriate sensory rewards will be needed to maximize efficacy and allow the development of a library of sensory rewards. We envision a "Level 1" option in the software that will allow the software to learn each client's individualized preferences for sensory rewards. A variety of alternative sensory rewards can be imagined, and are being explored for future development. Age-effectiveness of each type and category of reward will be explored, and individual preferences can be learned via a training session in which the app uses (and stores) the learner's preferred sensory reward.

Further capabilities such as the use of eye-tracking and facial behavioral analysis will be explored as input for evaluation of utterance effort (see Norbury et al., 2013, for example).

In the long term, dietary, supplements, and medical interventions may be studied.

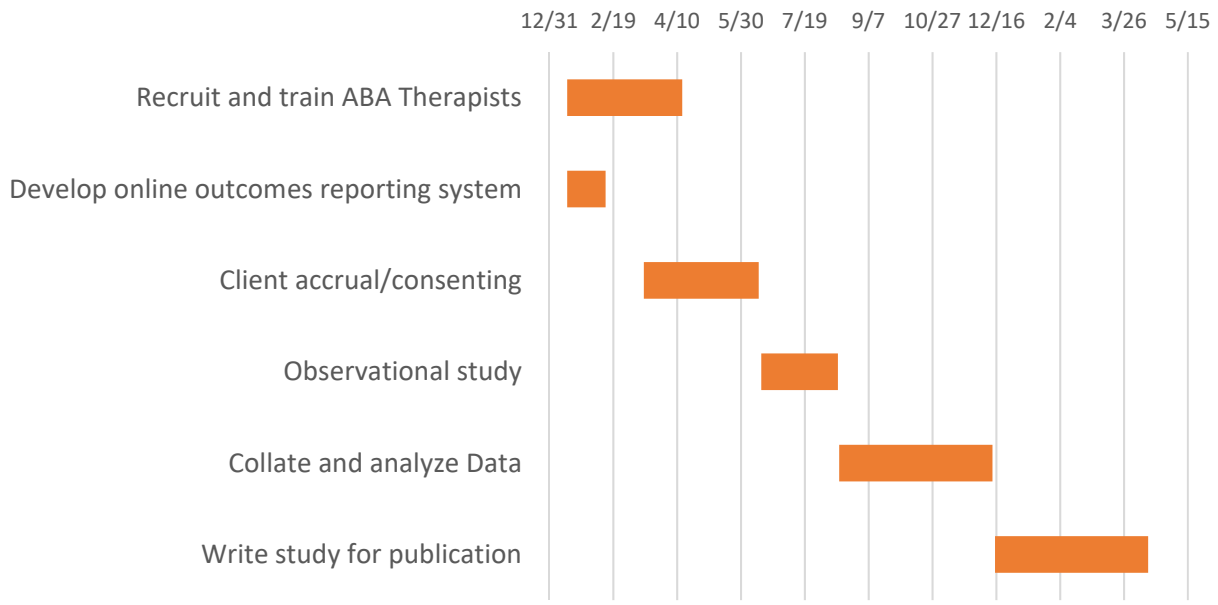
Added Value: The ASD Outcomes Research Competent ABA Team

Many studies from other investigators are being planned that will benefit from the establishment of an army of field-capable ABA Therapists trained to be competent in executing ASD baseline and outcomes research designed to assess the effects of emerging interventions and treatments for ASD. The value of this resource to the greater ASD research community cannot be underestimated. We can provide letters of support from researchers in ASD who would use this resource.

Timeline

Specific Aim 1: January 2018-April 2019

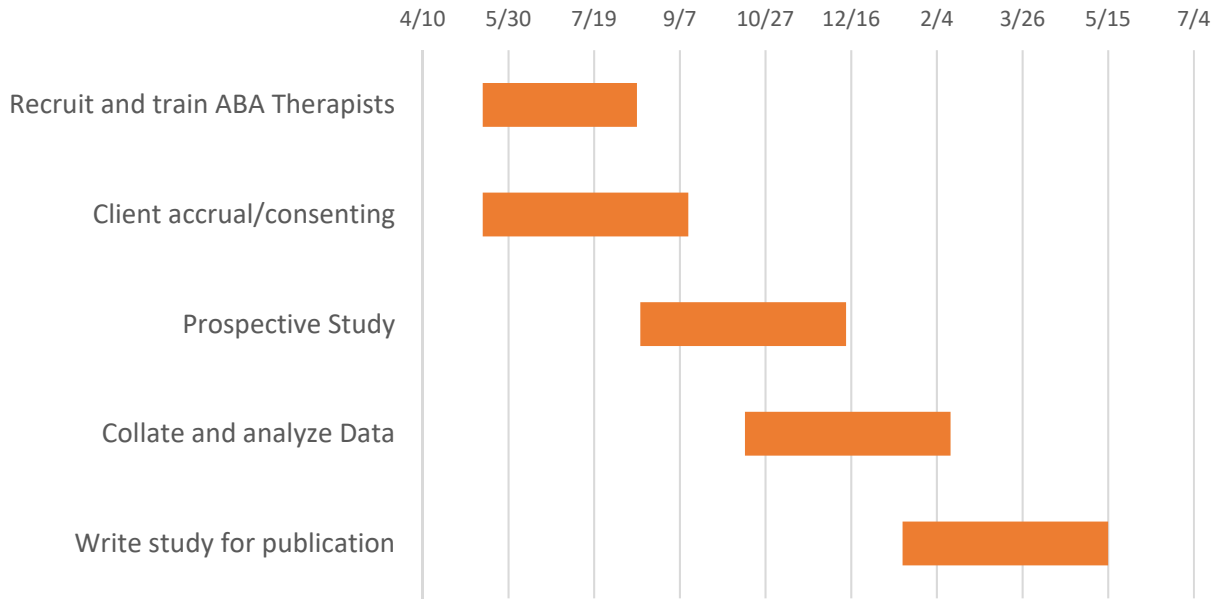
Tasks: Recruit and train ABA Therapists (3 mo), concurrently develop and standardize outcomes recording system (online) for language components of Vineland II, conduct observational study (2 mos), collate and analyze data (3 mos), write study for publication (3 mos).



GANT Chart, Specific Aim 1

Specific Aim 2: May, 2019-May 2020

Tasks: Recruit and Train ABA Therapists (3 mo), conduct randomized prospective trial (2 mos), update standardized outcomes recording system, collate and analyze data (3 mos), write study for publication (3-4 mos).



Specific Aim 3: Feb 2019-Dec 2019

Tasks: Recruit and train ABA Therapists, conduct randomized prospective trial (2 mos), collate and analyze data, write study for publication. Training will be conducted via webinars with online evaluation using test cases (with video). Qualified ABA Therapists will pass the evaluation with >95% proficiency within three attempts.

CCC-2 Source:

http://www.pearsonclinical.com/language/products/100000193/childrens-communication-checklist2-us-edition-ccc-2.html

Vineland 3 Source:

https://www.pearsonclinical.com/psychology/products/100001622/vineland-adaptive-behavior-scales-third-edition--vineland-3.html

Literature Cited

Ellis Weismer S, Kover ST. 2015. Preschool language variation, growth, and predictors in children on the autism spectrum. J Child Psychol Psychiatry. 56(12):1327-37. doi: 10.1111/jcpp.12406.

Kover, S et al., 2016. Ages of Language Milestones as Predictors of Developmental Trajectories in Young Children with Autism Spectrum Disorder. J Autism Dev Disord. 46(7): 2501–2507. doi:10.1007/s10803-016-2756-y.

Whitaker, HJ, 2005. Tutorial in Biostatistics: The self-controlled case series method. Statist. Med. 25(10):1768-97.

Mottron L et al., 2017. Should we change targets and methods of early intervention in autism, in favor of a strengths-based education? Eur Child Adolesc Psychiatry 26(7):815-825. doi: 10.1007/s00787-017-0955-5.

National Professional Development Center on Autism Spectrum Disorders. 2010. Pivotal Response Training Module. Evidence-Based Practice Brief: Pivotal Response Training. http://autismpdc.fpg.unc.edu/sites/autismpdc.fpg.unc.edu/files/imce/documents/PRT-Complete-10-2010.pdf

Norbury CF et al., 2013. Sources of variation in developmental language disorders: evidence from eye-tracking studies of sentence production. Philos Trans R Soc Lond B Biol Sci. 369(1634):20120393. doi: 10.1098/rstb.2012.0393.

Petersen I, Douglas I, Whitaker H. 2016. Self controlled case series methods: an alternative to standard epidemiological study designs. BMJ. 354:i4515. doi: 10.1136/bmj.i4515.

Rubin, Z. 2017. Development and Evaluation of Software Tools for Speech Therapy.

Rubin, Z., Kurniawan, S., Gotfrid, T., Pugliese, A. 2016.Motivating Individuals with Spastic Cerebral Palsy to Speak Using Mobile Speech Recognition. Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility: 325-326.

Rutter, M.; Le Couteur, A.; Lord, C. Autism diagnostic interview—revised. Los Angeles: Western Psychological Services; 2003.

Wodka EL, Mathy P, Kalb L 2013. Predictors of phrase and fluent speech in children with autism and severe language delay. Pediatrics. 2013 Apr;131(4):e1128-34. doi: 10.1542/peds.2012-2221. Epub 2013 Mar 4.

Wittke K, Mastergeorge AM, Ozonoff S, Rogers SJ, Naigles LR. 2017. Grammatical Language Impairment in Autism Spectrum Disorder: Exploring Language Phenotypes Beyond Standardized Testing. Front Psychol. 8:532. doi: 10.3389/fpsyg.2017.00532. eCollection 2017.

Verschuur, R et al., 2017. Increasing Opportunities for Question-Asking in School-Aged Children with Autism Spectrum Disorder: Effectiveness of Staff Training in Pivotal Response Treatment J Autism Dev Disord. 2017; 47(2): 490–505.

Appendix 1 AURS Study ABA Therapist Survey (Clinician End-User's Perspective)

Appendix 2 AURS Study Client Accrual Instructions – AURS Study

Appendix 3 AURS Study Informed Consent Instructions and Form

Appendix 4 IPAK ASD Outcomes Research Competency Training Outline

Appendix 5 AURS Study Statement of Objectivity and Non-Bias

9.0