COMIRB Protocol
COLORADO MULTIPLE INSTITUTIONAL REVIEW BOARD
CAMPUS BOX F-490; TELEPHONE: 303-724-1055
Fax: 303-724-0990

Protocol #: 19-1962

Version 14 October 2024

Project Title: Physiological mechanisms of action relating to immediate and long-term therapeutic horseback riding intervention effects in a psychiatric population of youth with autism spectrum disorder
Principal Investigator: Robin Gabriels, Psy.D.
Version Date: 29 December 2021

I. Hypotheses and Specific Aims:

Autism spectrum disorder (ASD) is prevalent (1 in 59)[1] and has high healthcare utilization costs.[2] A majority of youth with ASD have co-occurring psychiatric disorders[3,4], including higher than typical rates of depression, anxiety, and related aberrant behaviors (e.g., irritability, hyperactivity). Mental health treatment for this population is challenging due to their social, communication, and emotion regulation impairments that can collectively and detrimentally affect daily social functioning[5] and caregiver quality of life.[6-8] Mounting evidence suggests that emotion regulation deficits underlie aberrant behaviors in ASD[9], and that human-animal interactions can induce regulated/calm physiological states in humans,[10] including those with ASD.[11] This is perhaps one reason why an estimated 70% of families of youth with ASD seek alternative treatments, such as Therapeutic Horseback Riding (THR).[12] Our previous randomized controlled trial (RCT) established the efficacy of a 10-week THR group with ASD youth with aberrant behaviors.[13] Specifically, the THR group had medium effect size reductions in irritability and hyperactivity and significant improvements in social communication, social cognition, and word fluency compared to a no-horse barn activity (BA) control group.[13] Significant outcomes emerged by week five of the THR invention.[13] The BA control showed some significant within-group improvements, but no conclusions could be drawn due to the lack of a no-intervention (wait list) control. Exploratory analyses showed stronger positive outcomes in the subgroup of THR participants with co-occurring psychiatric disorders and initial outcome gains were maintained at six-month follow-up with a subset of THR participants.[14] Although promising, the physiological mechanisms relating to these outcomes remain unknown. It is also uncertain how long benefits persist compared to a no-intervention control, and whether ASD youth with co-occurring psychiatric disorders would respond similarly. **The goal of this project is to assess physiological mechanisms underlying THR's significant positive effects on ASD youth, particularly those with co-occurring psychiatric disorders, and evaluate durability, dose, and sub-population effects of the intervention.** Our long-range goal is to empirically establish THR as an ecologically valid, transdiagnostic intervention for youth with a variety of mental health issues, facilitating acquisition and maintenance of critical life skills that enhance quality of life (QOL) for individuals and their caregivers.

The objective of this proposed RCT is to identify physiological mediators of our well-tested[13,15] 10-week manualized[16] THR intervention compared to a similarly structured BA control group in up to 250 youth ages 6-16 yrs. with ASD and co-occurring psychiatric disorders recruited from the community, outpatient clinics, and a psychiatric hospital setting. An additional 20 ASD youth with co-occurring psychiatric disorders will be recruited for the exploratory aim's waitlist control and Hybrid intervention groups.

Specific Aim 1: Examine physiological mediators in a THR intervention group compared to a BA control on multivariate outcomes by measuring individual physiological response profiles and analyzing them for predictive validity and mediating effects on primary (irritability, hyperactivity), secondary (social communication, social cognition, and word fluency) and expanded outcomes (emotional dysregulation, caregiver QOL, and crisis mental health care usage). *Hypothesis 1a (H1a). THR group will have significantly greater improvement in outcomes collected at baseline compared to post-intervention. H1b: Post-intervention (week 10) afternoon salivary cortisol levels will be significantly reduced (indicating less stress) in the THR group compared to baseline/week one levels. H1c. THR group will have a decrease in heart rate and electrodermal activity and an increase in heart rate variability, both indicating a calm state, from baseline to post- intervention. H1d. Physiological measures collected at intervention midpoint and post-intervention) will be correlated with and largely explain/mediate the effects of THR on post-intervention outcomes.*

Specific Aim 2: Evaluate the durability of Aim 1 outcomes in the THR group compared to the BA group control six-months post intervention period. *H2. THR group will maintain significantly greater outcome gains six-months after the intervention compared to the BA control.*
Specific Aim 3: Explore dose and sub-population effects of THR and BA interventions by comparing effect size differences in THR and BA groups to a (1) 10-week wait-list control group (n = 20); (2) Hybrid intervention group (n = 20) (five weeks BA followed by five weeks THR); and (3) most severely impaired subsample (n = 20) of the THR study population randomized following psychiatric hospitalization. *H3a. The BA group will show significantly greater improvements on Aim 1 outcomes compared to those randomized to the waitlist control group. H3b. The Hybrid group will show a similar response pattern as the THR group.*
Specific Aim 4: Explore anecdotal observations of the study intervention personnel (i.e. instructors and volunteers) by providing study intervention personnel with a written survey of questions to gather information about their study observations and experiences. *H4. The written reflections of study personnel about their interaction with the study population will provide anecdotal support for the study findings.*

Our preliminary data demonstrate that physiological data collection is feasible with ASD youth engaged in THR at the proposed PATH international [17]premiere accredited riding centers: (1) northern Colorado where access to psychiatric care is limited and (2) Portland, ME near a specialized ASD psychiatric hospital program. We also partially replicated our previous results, demonstrating cortisol is a viable target mediator of THR effects on aberrant behaviors in ASD youth.[18] Our team has collaborated across the two study sites for over five years and has the necessary clinical and technical experience to conduct this next phase of THR research.

Background and Significance:

ASD population with co-occurring psychiatric disorders

Population overview: Autism Spectrum Disorder (ASD) is a critical public health concern due to increasingly high prevalence rates (1 in 59)[1], high healthcare utilization costs[2], and the fact that two-thirds of this population have co-occurring psychiatric disorders.[4] Compared to the general population, individuals with ASD are at significantly higher-risk for a range of co-occurring psychiatric symptoms and disorders, particularly anxiety, ADHD and mood disorders.[3,4,19,20] [21-23] Moreover, this population frequently engages in aberrant behaviors (e.g., irritability, hyperactivity, self-injury, aggression, etc.) that require intensive interventions.[20,24,25] that present a major source of caregiver stress and burden[7,8,26], amplify social and communication deficits, intensify restricted, repetitive and stereotyped behaviors[27], and complicate the ability to engage safely in community settings. Aberrant behaviors require high levels of intensive interventions (e.g., special school placements, psychopharmacology, in-home therapies, and psychiatric hospitalization), often beyond what providers in community settings can manage.[20,24,25] Despite the high prevalence of psychiatric disorders and related aberrant behaviors in ASD, interventions, including human-animal interventions (HAI), targeting this psychiatric subpopulation have been understudied.[28]

Youth with ASD and psychiatric diagnoses seek treatment in psychiatric hospital settings at much higher rates than non-ASD youth.[29] Examining an ASD psychiatric cohort (n = 350) from our multi-site study (Autism Inpatient Collection; AIC) of youth (ages 4-20 yrs.) admitted to six specialized psychiatric inpatient units revealed they had dangerously high levels of irritability behaviors (e.g., physical aggression, severe tantrums, heightened reactivity, self-injury) at admission.[30] Such irritability levels measured by the Aberrant Behavior Checklist-Community (ABC-C)[31] Irritability score (mean = 29.7)[30] were much higher than the typical clinical threshold of ABC-C irritability levels (i.e., ≥14-16) in ASD psychopharmacology clinical trials.[32,33] Additionally, youth with ASD often have difficulty self-reporting internal experiences associated with emotions and distress.[34] As psychiatric service demand for this subpopulation of youth with ASD increases, more healthcare dollars are spent.[35] It is thus imperative to evaluate safe and effective treatments that target underlying mechanisms leading to clinically meaningful and long-term change in this population.

Emotion dysregulation and physiological response patterns in ASD: Emotion regulation is the process of organizing emotional responses via autonomic, cognitive, and behavioral modification to address environmental demands.[36] Mounting evidence suggests that emotion dysregulation (ED) contributes to a variety of psychiatric disorders,[37] has a multitude of negative downstream effects[38,39], and is a common problem in ASD[9]. Manifestations of ED in youth with ASD include irritability, hyperactivity, tantrums, aggression, elopement, property destruction, and self-injury.[9,24] ED is also thought to contribute to co-occurring psychiatric disorders (e.g., anxiety and depression) in the ASD population by prolonging and heightening distress levels.[37,44]

Measuring physiological responses (e.g., cardiovascular, electrodermal, cortisol) during states of heightened or experimentally manipulated emotion is a common research strategy in psychophysiology,[40] and growing evidence suggests that ED in ASD may be related to altered physiological arousal responses. For instance, individuals with ASD consistently show elevated norepinephrine levels[41-46], atypical cardiovascular baselines [63,67] and reactivity,[47] a lack of task-to-task variability,[48] increased behavioral/physiological hypersensitivity,[49,50] atypical responses to induced stressors[51-

[55], and larger baseline pupil dilation.[56]; all congruent with ED. Furthermore, reviews of cortisol in the ASD population report higher reactivity to daily stressors compared to TD children,[57,58] sluggish response patterns to stressors,[59] and altered responses in those with ASD who endorse having high levels of irritability.[60] Taken together. while significant heterogeneity exists in autonomic and hormonal responsivity, it is clear that physiological dysregulation can occur in ASD[61] and that physiological response patterns may underlie and be influenced by aberrant behaviors.

ASD Interventions: Treatment of individuals with ASD and co-occurring psychiatric disorders often requires psychopharmacologic and psychotherapeutic interventions to address the aberrant behaviors. Psychopharmacologic medications are frequently used to treat a variety of symptoms in this population.[62] Risperidone and Aripiprazole are the two antipsychotic medications that have empirically-supported evidence[63] and are approved by the U.S. Food and Drug Administration for significantly reducing symptoms of irritability in the ASD youth population, although drawbacks include medication side-effects. It has been proposed that THR might be a safe adjunctive intervention to facilitate lowering medication dosages in this population.[64] In our prior RCT of THR in youth with ASD[13] we observed reduced ABC-C Irritability scores approximating 50% of the difference in efficacy between low and high dosages of Risperidone.[65] Cognitive-behavior therapy (CBT) and behavioral interventions are also often employed to target aberrant behaviors in the ASD population. There is growing evidence that CBT can reduce anxiety symptoms in higher functioning ASD youth who have language abilities and higher intelligence levels (IQ≥70).[28,66] An off-shoot of CBT, mindfulness training (e.g., meditation and breathing techniques), has some preliminary beneficial evidence in higher functioning youth with ASD with increasing patient-reported QOL and decreasing repetitive thoughts.[67] However, there is limited evidence for the application of CBT to target externalizing symptoms of irritability/aggression in ASD youth, particularly in those more severely affected. Behavioral interventions, including Applied Behavior Analysis (ABA), have a large body of empirically-supported evidence for addressing a broad-range of adaptive skill deficits and aberrant behaviors in youth with ASD, including those with intellectual disabilities (ID).[68,69] A review of 101 studies of specific forms of behavioral interventions targeting externalizing behavior problems (e.g., aggression, self-injury, property destruction) revealed that a majority of study participants were treated in hospital or residential settings due to their severe symptom presentation and treatments resulted in a more than 80% reduction of problem behaviors compared to baseline.[24] However, drawbacks of these behavioral interventions are that they tend to require access to highly specialized clinician expertise, substantial time and intensity to see improvements that are not consistently attained or maintained,[28] and involve strategies difficult to implement in community settings.[24] Moreover, despite evidence for CBT and other behavioral interventions for ASD youth, most studies have not evaluated long-term effects. The extant body of ASD intervention research discussed by Lerner et al., has "…barely begun to explore the common unique processes by which these interventions 'work', the conditions under which they 'work best' and for whom which type of treatment might be optimal".[70] The significance of the current proposal includes our aims to assess the putative causal mechanisms, durability and dose of THR for the majority psychiatric subpopulation within ASD, for whom our preliminary evidence indicates may be best responders. Our proposal has the potential to provide evidence for THR as an important adjunct to "mainstream" ASD interventions, targeting a wider functioning range of this population.

Premise for the proposed project

HAI Research with ASD: The HAI field has focused thus far on measuring the efficacy of these interventions, with limited focus on providing evidence to explain the historical theory that animals elicit beneficial arousal responses in humans. A 2017 systematic review of 28 studies suggests that HAI can collectively improve sociability, communication, positive emotions, and reduced arousal levels in youth with ASD.[71] That same review highlights that HAI research with equines was one of the most common methods employed, wherein dosing of HAI was typically 8-12 weeks.[71] A systematic mapping review of studies involving equine-assisted activities and therapies (EAAT) with the ASD population revealed a wide variety of intervention methods and treatment targets. [72] They range from equine assisted activities (EAA) (e.g., psychoeducational horseback riding, therapeutic riding) involving riding instructors, coaches or trainers) to equine-assisted therapies (EATs) (e.g., hippotherapy, simulated developmental horse-riding) involving therapists (e.g., occupational or physical therapists) and therapeutic riding instructors.[72] Outcome improvement areas reported by EAA studies included social interactions, communication, sensory processing, movement control, ASD symptom severity, and QOL, whereas EAT studies reported outcome improvements in motor control and adaptive living skills.[72] While there is limited research and conflicting results regarding the long-term effects of EAATs,[73] our six-month pilot follow-up of THR provides the first known prospective study to demonstrate maintenance of initial outcome gains compared to a control.[74] A number of systematic reviews have highlighted multiple concerns about the lack of methodological rigor in EAAT and HAI research in general, which limits the ability to demonstrate its efficacy, and very few have attempted to investigate biological mechanisms that may explain reported effects.[71,72,73,75,76]

Some studies have assessed physiological response patterns (e.g., cortisol, cardiovascular, electrodermal) in individuals engaged in HAI and concluded that HAI may produce both a calming or regulated [10] as well as an alert and motivated state of arousal. [77] Specifically, Beetz et al., (2012)[10] summarized results from six HAI studies showing a

decrease in salivary cortisol levels in response to interactions with dogs. One of these studies examined the effect of service dogs on 42 youth with ASD and findings included statistically diminished cortisol awakening response (CAR) levels (i.e., 30 minutes after waking) when service dogs were introduced.[78] A few studies have examined changes in cortisol levels in response to equine interventions. For instance, an RCT with TD youth, participating in an 11-week equine-facilitated learning program, found lower basal salivary cortisol levels compared to controls.[79] A study of hippotherapy (n=8, male, ASD) resulted in significant salivary cortisol reductions after versus before riding.[80] Several HAI studies with adult and child populations have observed lowered blood pressure, heart rate (HR), and skin temperature along with increased heart rate variability (HRV) in response to both familiar (pets) and unfamiliar animals, suggesting stress response reduction.[10] Conversely, a study of 36 youth engaged in a reading activity found increased arousal levels (higher mean HR and area under the curve salivary cortisol values) when a dog was present versus absent.[77] It is important to note that these physiological measurements are taken either before/after HAI intervention or during low activity (i.e., interacting with a service dog). However, during THR, a physical activity requiring higher than baseline autonomic output[81], we would expect a transient *increase* in arousal due to the aerobic demand, and subsequent arousal reduction during both pre- and post-riding baseline over time. This is the pattern commonly seen in exercise interventions,[82,83] including those used for stress reduction.[84] In summary, HAI appears to have to a direct effect on physiology and can be successfully measured in humans, including youth with ASD. However, the physiological mechanisms associated with HAI outcome gains is in its infancy. Further rigorous examination is required to advance our understanding of individual differences in physiological responses to HAI, including the degree to which short (i.e. within session) versus longer-term (e.g., across session) benefits are observed in ASD youth.[18]

Physiological arousal mechanisms underlying THR for ASD youth with co-occurring psychiatric disorders. Based on extant literature and our previous THR research with the ASD population, our guiding theoretical framework is that THR promotes positive behavioral and emotional health outcomes mediated in part by inducing more regulated physiological arousal states over time. This framework is consistent with positive training effects observed in physical activity interventions; the concept of mindfulness[85], and the curvilinear relationship between physiological arousal and performance demonstrated by the Yerkes-Dodson Curve.[86] Thus, we hypothesize that our manualized THR approach regulates/conditions physiological arousal levels, and is at least partially responsible for our previous observations of THR outcomes.[13] In an effort to directly test this hypothesis, we aim to gather several physiological indices including cardiovascular, electrodermal, and salivary cortisol levels, while controlling for commonly cited social motivation mechanisms of change in general psychosocial research (i.e., therapeutic relationship, group membership, social knowledge) through use of an RCT design.[70] **(See Fig. 5 in Analysis section for the model to be tested)**.

In summary, there has been a call to examine mechanisms to test the long-standing theory that interacting with animals can improve mental health wellness in humans.[87] Studying mechanisms of change must include a theoretical plausible and testable explanation.[88] There are a variety of empirically validated ASD-intervention models, yet there has been limited focus on designing ASD intervention trials to address mechanisms of change [89] to help explain how ASD interventions work to achieve positive treatment response.[90] These are particularly critical variables to research for the majority of the ASD population with co-occurring psychiatric disorders, given their related aberrant and challenging behaviors, high-need for intensive intervention services and variable outcomes.[89,91] Improved understanding on how to tailor treatments to those who may respond best to specific interventions can lead to cost-effective and efficacious intervention implementation.[70,92] Previous research, combined with our preliminary data, suggest there is high potential value in targeting THR for the psychiatric ASD population and that physiological activity is a putative mechanism of therapeutic gains that warrants further study.

Our project design is an innovative approach that will provide ground-breaking advancements in the HAI field (specifically THR). Our approach innovates by: conducting a large RCT targeting a majority subset of the ASD population, who have indications of being potentially stronger THR responders[99]; using an intervention manual for both experimental and control conditions and measuring their fidelity to ensure implementation uniformity, specifically accounting for effects of the horse while controlling for social attention throughout the study; including a method to understand the six-month maintenance of THR outcomes vs. a control; and implementing the study in two community-based rural settings to expand the ecological validity of our manual-based THR intervention.[16]

We are grounding our mediation analysis in the framework of a model to be tested using a RCT. Methodologically, including multiple physiological measures will allow us to decrease the biasing effects of measurement error[58] and more systematically assess whether positive THR outcomes are partially accounted for by physiological response profiles over time in youth with ASD and co-occurring psychiatric diagnoses.

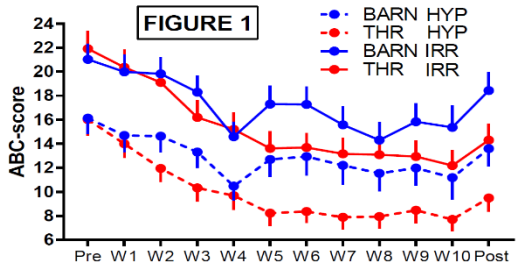
Our measurement approach is technically innovative. Our proposal seeks to employ objective measures of participants' experiences by using unobtrusive state-of-the art, ambulatory physiological measures during an HAI activity to capture the potentially complex profile of mediating factors that may explain previous outcomes. Our methodology

takes advantage of technological advances in wearable physiological biosensors that overcome the inherent difficulty obtaining reliable self-reports on emotional and physiological states in individuals with ASD, a portion of whom are either minimally-verbal or have alexithymia.[103-105]

.

Preliminary Studies/Progress Report:

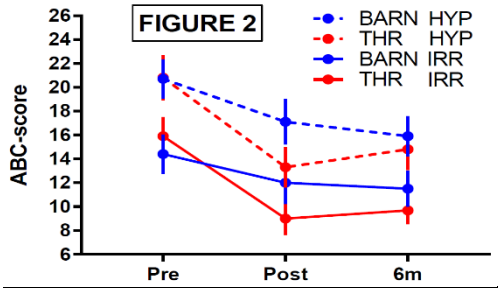
Preliminary data overview: Frequently, the HAI field is criticized for using inconsistent research methods, as this limits the field's ability to establish empirical support. Our THR intervention program of research is methodologically consistent and follows guidelines for research to validate ASD interventions that include: (1) refining intervention methods and conducting initial efficacy studies; (2) developing intervention and comparison control manuals and evaluating intervention outcomes at different sites; (3) conducting RCTs and evaluating mediator and moderator variables; and (4) assessing community providers' intervention implementation fidelity.[97] Our HAI research began with an **initial pilot study**[15] to evaluate the efficacy of our 10-week THR intervention with ASD youth (n = 41) ages 6-16 yrs. That pilot provided preliminary evidence for our THR intervention enabling significant improvements in aberrant behaviors (irritability, lethargy, stereotypic behaviors, and hyperactivity) in the THR group compared to a waitlist control.[15] Our follow-up THR studies include a RCT evaluation of our refined THR manual approach[16] compared to a no-horse BA control with youth diagnosed with ASD, demonstration of our THR and BA implementation fidelity at two different riding centers,[18] a pilot analysis of the six-month maintenance of THR outcomes gains,[74] and two pilot studies demonstrating the feasibility of collecting the physiological measures to be utilized in the current proposal. **RCT of THR.**[13] Our large (N = 127) RCT established the efficacy of a 10-week, one-hour THR group for youth (ages 6-16 yrs; IQ ≥ 40) with ASD and high ABC-C[98] baseline mean irritability scores (THR mean = 16.0; control mean = 16.1).[13] Compared to a similarly structured BA control group, the THR group had significant improvements in behaviors, social skills, and word fluency.[13](**See Table 1** for specific areas of improvement and effect sizes; ES). The groups did not differ significantly on any demographic variables (e.g., IQ, gender, community psychiatric diagnoses, use of psychotropic medications). Caregiver-reported psychiatric disorders for this sample were mood, anxiety, and ADHD, and these psychiatric disorders were present in 48% of participants in both THR and BA groups. Significant between-group reductions in irritability and hyperactivity levels (compared to baseline) emerged as early as the fifth week of the THR invention (**See Fig. 1**).[13]



Results suggest that week five of the intervention may be an optimal time to evaluate mediator variables (e.g., physiological) that may influence outcomes and that five weeks of THR may be an optimal dose to effect positive changes.

Exploratory analyses of the RCT data[13] showed stronger positive outcomes in the subgroup of THR participants with co-occurring psychiatric disorders compared to those without (**See Table 1**). ***Conclusions:*** *Results support targeting the psychiatric subpopulation who arguably stand to benefit the most from THR and have a higher likelihood of showing significant THR effects, a prerequisite for mediation analysis. The Interquartile range of the ABC-C irritability score for this subgroup was 8 (25th percentile), which will be used as the minimum inclusion criteria for this study. This inclusion cut-off is much lower than that of medication trials (e.g., 18),[33] which allows our study to target a wider range of mild to severe symptom presentations.*

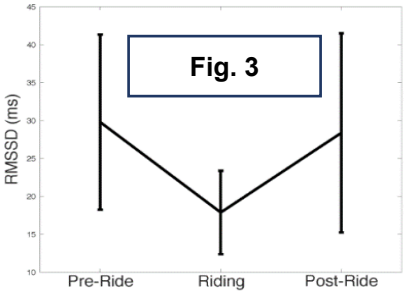
THR six-month outcomes pilot.[14] Our six-month follow-up study of a subset of participants from the RCT[13] (THR n = 36; BA control n = 28) revealed that the THR group maintained their reductions in irritability behavior (effect size=0.32), but not in hyperactivity compared to the BA control. **(See Fig. 2)**. We only examined the THR group for other outcome measures, which indicated sustained significant improvements in social communication and word fluency. ***Conclusions:*** *The small sample size and the fact that we did not assess efficacy of THR compared to the BA control group on all outcome measures employed in the RCT[13] limits the validity of these pilot study results and warrants a more thorough evaluation of whether the effects of THR can be maintained for at least six-months after the intervention compared to a control. For the current proposal, we anticipate lower attrition than the 20% rate in this pilot for the following reasons: (1) outcome evaluations will occur in closer proximity to participants' residences; (2) increased monetary incentives will be provided for the 6-month follow-up; and (3) the BA control group will receive one free riding lesson after their 10-week BA intervention.*

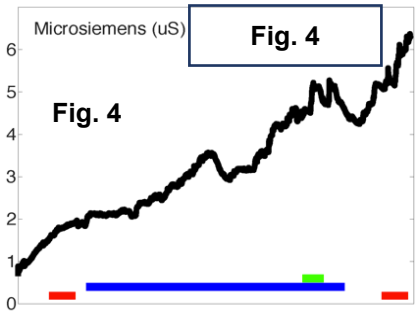


THR replication trial and salivary cortisol sampling.[99] This pilot aimed to replicate our RCT[13] at a novel riding center (i.e., Hearts & Horses, one of the centers proposed in this project) and evaluate THR effects on salivary cortisol levels in 16 youth with ASD (ages 6-16 yrs). Of note, 94% of this sample had community-based psychiatric diagnoses and a mean baseline ABC-C[98] irritability score of 17. Participants were randomized to either a THR group (n = 8) or BA control group (n = 8) stratified by NVIQ (≤85 or >85). Salivary cortisol samples were collected weekly at a consistent afternoon time immediately before and 20 minutes after intervention conditions. Intent-to-treat analysis (n = 16) revealed results similar to our RCT.[13] Compared to the control, THR participants had large Cohen's d improvements in hyperactivity (ES =1.49), social awareness (ES=1.54), irritability (ES=1.08) and social communication (ES=1.17) behaviors. There were no significant improvements in number of words or new words spoken during the standard language sample and no significant THR intervention effects on cortisol. ***Conclusions:*** *Although the small sample size may have contributed to the lack of cortisol results, this study demonstrates our ability to collect successfully salivary cortisol with this population within the riding center environment. This is the first known study to report partial replication of results from a previous RCT of THR. This study of ASD youth with psychiatric disorders also reinforces the hypothesis that this subpopulation may benefit more from THR, making it more likely to detect mediator effects in this subpopulation.*

Ambulatory cardiovascular response patterns to a canine-assisted activity in psychiatrically-hospitalized and diagnosed ASD youth. This pilot study investigated participants (ages 6-18yrs) with a mean ABC-C irritability score of 26.5 (SD=8.3), measuring their peripheral physiological responses via the wrist-worn E4 device [100] to a 10-minute canine and handler treatment condition (DOG) compared to a 10-minute novel toy and handler control condition (TOY)[101]. In both conditions, HR slowed overall (η^2=0.237; DOG, d=0.37, p=0.046; TOY, d=0.64, p=0.002) and HRV increased overall (η^2=0.465, DOG, d=0.38, p=0.08; TOY, d=0.64, p=0.003), consistent with increased cardiac parasympathetic modulation (i.e., reduced stress response). However, no interaction was observed (p=0.23, p=0.75) with improvements equivalent in both conditions. While n=41 participants returned data, repeated measures ANOVA requires four individual values per person, reducing the sample to n=15 observations overall (paired comparisons n=23 and n=25). ***Conclusion****: This pilot demonstrates the feasibility of collecting biosensor data in a similar sample to the proposed study, and the ability to observe reactivity in animal and non-animal conditions. This pilot highlights need for more robust methods of gathering peripheral physiological data and increasing sample size to achieve appropriate power to account for potential data loss, both of which are accounted for in the current proposal.*

Feasibility pilot of ambulatory cardiovascular and electrodermal activity monitoring in ASD youth during THR. We assessed the feasibility of ambulatory peripheral physiological data collection in ASD youth while engaged in THR lessons over five contiguous weeks at the two centers (Colorado & Maine) proposed in this project. Eight ASD youth (6-13 yrs.; mean age = 9.1 yrs; mean ABC-C irritability score=15.8 (SD=11.10); 75% with co-occurring psychiatric diagnosis, including 38% with multiple) wore the Empatica E4 with two electrodes attached to the underside of the wrist of their non-dominant hand to record electrodermal activity (EDA) and the Faros eMotion 180 with three electrodes attached to the thorax in standard configuration to record HR and HRV. **Fig. 3** shows the ensemble average of all participants for time domain HRV (RMSSD) pre-baseline, the riding section, and post-baseline. As expected, the riding section demonstrably reduces HRV. As individual differences in eccrine sweat gland density prevent ensemble averaging of EDA (means ranged from 0.5uS to 30uS); a single participant (Maine, Green, 1) over the entire period of collection is shown in **Fig. 4**. Baseline sections (red) show noticeably less EDA variability than riding (blue); fast riding (green) shows immediate transient increases. Moreover, across all cardiovascular and electrodermal activity data collection sessions, only 15% were lost due to equipment issues, non-cooperativeness, or environmental noise related to equine movement. ***Conclusions:*** *ASD youth with psychiatric diagnoses tolerate wearing ambulatory peripheral physiological devices, they do not interfere with THR, they produce quality signals sufficient for analysis, and indicate sensitivity to non-riding and riding conditions. Based on these pilot results, we estimate that future data retention will be approximately 85%, consistent with pediatric studies in the typically developing population.*[102]





Research Methods

Study Measures. The following descriptions are for those measures to be used for screening, demographics, mediators, or outcomes that have not been previously described (i.e., ADOS, Leiter, SCQ) by this research team.[13,15] **See**

Appendix for table of study measures informants, and collection timelines along with E-4 and Faros 180 equipment application procedures.

Child and Caregiver Information Form (CCIF)[113] is a caregiver-completed form that provides relevant study demographics, including participants' current medical and psychosocial treatment information. The CCIF has been modified to include information for calculating socioeconomic status (SES).[114]

Child & Adolescent Symptom Inventory 5 (CASI-5)[107] is caregiver-report measure used to screen for psychiatric symptoms in children and adolescents ages 5 to 18 y/o. The CASI-5 assesses for symptoms of a variety of disorders found in the DSM-5.[27] CASI-5 subscale scores show a high degree of predictive and concurrent validity[115,116] and are demonstrated to have internal construct validity when evaluating psychiatric concerns in youth with ASD who range in verbal ability.[117,118] The CASI-5 yields **Symptom Count score** (number of clinically concerning symptoms) and a **Symptom Criterion score** (number of symptoms needed for a specific DSM-V[27] diagnosis). **Participants' Symptom Count score must be greater than or equal to the Symptom Criterion score to meet the Screening Cutoff score for a particular DSM-5[27] diagnosis.** The following symptom subscales will be used for this study: **Attention Deficit Hyperactivity Disorder (inattentive, hyperactive-impulsive, & combined types), anxiety disorders (generalized anxiety, social anxiety, separation anxiety obsessive-compulsive, and related disorders, including trauma and stressor related conditions), and mood disorders (major depressive episode, dysthymia, manic episode, disruptive mood dysregulation**).

Salivary Cortisol: Cortisol can be measured accurately in a minimally invasive manner via saliva, and the hypothalamic Pituitary Adrenal (HPA) axis is fast acting in response to environmental conditions. Typically, HPA activity shows a distinct non-linear circadian rhythm. In TD individuals, cortisol levels increase before awakening, show a rapid rise from awakening to +30-45 min, and a daily fall, with levels lowest during the normal sleep cycle before repeating.[119] Circadian cortisol rhythms in children with ASD are not always consistent with matched TD controls, instead displaying lower AM and higher PM levels.[120] However, preliminary evidence suggests that afternoon salivary cortisol levels in ASD youth are less irregular, indicating that afternoon sampling may be a more reliable measure of stress for this population.[121] Each participant will donate approximately 0.5-1 mL of saliva at the riding center immediately before and 20 minutes after each THR or BA group session weeks 1-10. Salivary cortisol collections will only be analyzed for weeks 1, week 5 or the adjacent week if participant misses session 5, and week 10. Collection methods will be the same as previously reported in our replication pilot.[OBJ] In brief, THR and BA interventions will occur in the afternoons (between 1:00-5:00 PM) to avoid the sharp rise and fall in cortisol levels in the AM associated with diurnal patterns of HPA axis activity. Only saliva samples taken at weeks 1, 5, and 10 of the intervention will be assayed for cortisol. Saliva samples will be collected using absorbent swabs designed for use with youth (Salivabio, LLC, Carlsbad, CA). They will be temporarily stored frozen at the riding centers, then transported to the respective hospital research site for storage in a freezer until shipped to Salimetrics for analysis. Caregivers will receive the following instructions prior to collection days: (1) Participants should avoid eating a substantial meal (other than a snack) and drinking liquids other than water within 30 minutes prior to collection time and (2) cannot brush their teeth for at least 45 minutes prior to collection time. Participants who do eat within the hour before collection will be asked to drink water and then wait 10 minutes before collections.[111]

Electrodermal Activity (EDA): The Shimmer3 EDA has been successfully used in studies worldwide involving youth with ASD similar to our target population, and is the frontline instrument used in the lab of our physiology expert co-investigators. Continuous peripheral physiological arousal will be assessed in all participants using Shimmer3 EDA monitor (also known as skin conductance or galvanic skin response). This is a well-established index of sympathetic nervous system arousal.[40] The Shimmer3 EDA weighs 28 grams (dimensions 3 x1 ½ x ½ inches), is made of durable plastic materials that make it water resistant and shock proof, and is regulatory compliant. The Shimmer3 EDA monitor records EDA from the ventral area of the wrist using alternating current imperceptibly applied to the skin through two hypoallergenic, durable, and replaceable Ag electrodes. As dry non-adhesive electrodes tend to move on the skin surface during physical activity and destroy signal quality, disposable adhesive electrodes with an appropriate salinity (e.g., 0.5%; EL507) will be used, as biopotential electrodes with higher salinity overestimate tonic skin conductance. EDA sampling frequency in the Shimmer3 is 4 Hz with a 0.01– 100uS range. Peripheral skin temperature is recorded by the Shimmer3 EDA monitor at 4 Hz in the -40 to 115-Celsius range using optical infrared thermopile. Finally, the Shimmer3 EDA records motion-based activity up to ±8g at 32 Hz using 3-axis accelerometry. EDA recordings from the central wrist are well correlated (r =.574) with traditional laboratory-based measurements from the fingers.[122] Although the Shimmer3 EDA assessments will be taken weekly during the 10-week intervention, these assessment data points will only be analyzed for weeks 1, week 5 or the adjacent week if participant misses session 5, and week 10.

Heart Rate (HR) and Heart Rate Variability (HRV): The Shimmer3 cardiac (ECG) monitor has also been successfully used in studies involving youth with ASD similar to our target population, and is the frontline instrument used in the lab of our physiology expert co-investigators to measure participants' HR and HRV. The Shimmer3 ECG monitor weighs 28 grams (dimensions 3 x1 ½ x ½ inches). This cardiovascular monitor was specifically developed to record electrocardiographic signals while participants are active or in motion. The form factor is compact (about 2" by 1"), lightweight (13g), and designed to wear continuously and unobtrusively. It samples ECG and HRV up to 2,048 Hz, which is sufficient for research.[125] It has a built-in accelerometer for quantifying periods of physical activity, necessary to identify

signal artifacts in post-processing. Although the Shimmer3 ECG assessments will be taken weekly during the 10-week intervention, these assessment data points will only be analyzed for weeks 1, week 5 or the adjacent week if participant misses session 5, and week 10.

Physiological Data Interval Recording Log: During weeks *1-10* of the THR and BA interventions, study personnel will record the start and end times using a stopwatch synchronized to system clocks in the Shimmer3 EDA and ECG monitors for each of the following activity periods: 3 to 5-minute baseline (art table activity); 45-minute THR (mounted) or BA lesson; 15-minute THR (grooming) or BA (scrapbook) closure activity; 20-minutes post (art table activity). These records will enable precise segmentation of cardiovascular and electrodermal activity data, enabling cross-condition analyses within-session and across-session change over time.

Physiological Sample Collection Log: During weeks *1-10* of the THR and BA interventions, caregivers will complete a log to report information that may affect physiological measurements, including participant's sleep the night prior to the lesson,[41,42] particular stressors that occurred that day, and medications or physical health status changes.

Aberrant Behavior Checklist-Community (ABC-C)[98], a 58-item behavior symptom checklist commonly used as a primary outcome measure in intervention studies with the ASD population, [98] is described previously.[13,15] The ABC-C will be completed weekly during the 10-week intervention by a designated consistent caregiver for each participant. The ABC-C Irritability scale will be used as a study screener (see inclusion criteria and screening methods sections.

Social Responsiveness Scale™, Second Edition (SRS™-2)[126] is a 65-item caregiver-report measure that evaluates social impairments in ASD. The SRS™-2 generates five treatment subscales (social awareness, social cognition, social communication, social motivation, autistic mannerisms), also described previously. [13,15]

Systematic Analysis of Language Transcripts (SALT)[127] provides standardized guidelines to elicit, transcribe, and analyze language samples from individuals, including those with ASD. The SALT also provides language analysis programs to compute a vocabulary diversity quotient from transcripts entered into the database. A five-minute expressive language sample will be elicited by each participant and recorded by the project's Speech Therapist, blind to participants' randomized group assignment, also described previously. [13,15] The speech therapists from Colorado and Maine sites will follow a standard administration protocol of prompts to obtain the 5-minute speech sample from participants at the pre- and post-intervention and 6-month time periods. The Colorado and Maine site speech therapists will develop administration and scoring reliability on the SALT[127] by the Colorado SALT trainer speech therapist reviewing a videotaped administration of the Maine site speech therapists SALT administration and then giving them feedback. The Maine and Colorado site speech therapists will also achieve an (at least) 80% scoring reliability on three consecutive SALT[127] transcriptions assessments.

Emotion Dysregulation Inventory (EDI)[128,129] The EDI was developed to assess ER impairment in ASD following guidelines from the NIH Patient-Reported Outcomes Measurement Information Systems (PROMIS) initiative, and its final items were based on factor analyses and item response theory (IRT) analyses using data from 1,755 youth with ASD. The EDI is a 30-item caregiver-report form that includes a 24-item Reactivity item bank, which captures intense, rapidly escalating, sustained, and poorly regulated negative emotional reactions, and a six-item Dysphoria scale characterized by minimal positive affect and motivation, and the presence of nervousness and sadness. Both the Dysphoria Scale and a seven-item Reactivity Short Form have IRT-based theta scores, which have a mean of 0 and SD of 1, and provide superior discriminative ability to raw scores.[130] Validity evidence includes expected group differences (higher scores in an ASD psychiatric inpatient versus community sample), correlations with measures of related constructs, and demonstration of test-retest stability.

World Health Organization's Quality of Life Instrument (WHOQOL-BREF)[95] is a widely used measure of patients' and caregivers' QOL. It has cross-cultural validity and strong psychometric properties (good discriminant, content, and test-retest validity). It is a 26-item measure that assesses the impact of the child on the caregiver's QOL in four domains (physical, psychological, social, and environmental). Test-retest reliability is good for the four domains (Cronbach alpha 0.66-0.84).[95] This tool was validated in parents of children with ASD and shows promise as being sensitive to unique challenges faced by parents of children with ASD.[94]

Crisis Mental Health Care Usage Survey is a list of questions developed by the Autism Developmental Disabilities Research Collaborative (ADDIRC) pertaining to the participant's need for crisis care services within 2 ½ months before starting the study, at the end of the10-week study intervention period, and 6 months after the study intervention period. Survey questions include psychiatric hospitalization (inpatient or partial), in-home crisis evaluation/management by police, behavior specialist or social services, or mental health related emergency room visits.

Fidelity Rating Tool[16] was developed for, used, and discussed in our prior THR studies.[13,15] It will be used to rate 20% of each 10-week intervention session (THR, BA and Hybrid). The THR & BA control group Fidelity Coordinators will rate video-recorded THR and BA sessions and achieve inner-rater reliability of at least 80% after observing two THR and two BA group lessons.

Intervention Team Reflection Survey is a set of four questions developed to capture anecdotal observations about the study participants from the intervention team personnel. The survey includes location of the team member, role within the study, and sessions team personnel participated in. The survey questions cover their overall experience, inspirational moments, observed changes in participants, and suggestions for improvements in future research.

Description of Population to be Enrolled:

Population Description: We will recruit up to 250 youth (ages 6-16 yrs) with a documented diagnosis of ASD and a co-occurring psychiatric disorder. Note: participants must be age 6-16 years to complete the 10-week intervention and all post intervention assessments, but may turn 17 years of age by the 6-month follow-up assessment. Participants must meet the following:

Inclusion criteria: (1) Aberrant Behavior Checklist-Community (ABC-C) [31,98] Irritability subscale score ≥8; (2) Leiter-III[106] Nonverbal IQ standard score ≥ 40; (3), meeting the Symptom Criterion score (minimum number of symptoms necessary for a DSM-V[27] mood, anxiety, or ADHD diagnosis) on the Child & Adolescent Symptom Inventory 5 (CASI-5);[107] and (4) a confirmed diagnosis of ASD by meeting clinical cut-offs on the Social Communication Questionnaire (SCQ)[108] (≥ 11)[109] and Autism Diagnostic Observation Schedule-2 (ADOS-2).[110] Only one child with ASD per family will be included in the study to maintain independent observation of participants. Participants will also require a consistent caregiver (i.e., parent or legal guardian) to complete study outcome measures. We will not exclude participants taking medications, as this is a common occurrence for the target population. However, information regarding medications, dosages, and any changes during the study will be collected from caregivers. Medications or substances that could affect cardiovascular or electrodermal activity (e.g., amphetamines, anxiolytics, anticholinergics, beta blockers) or cortisol levels.[111] (i.e., smoking or regularly use oral, inhaled, or topical steroids), will be flagged and participant's physiological data collected may be excluded or controlled for in the analyses.
Exclusion criteria: Potential participants will be excluded for any of the following reasons: (1) medical or behavioral issues (e.g., animal abuse or phobia to horses) that prevent participation; (2) being a ward of the state; (3) is judged during the riding center screening to have significant riding experience, as the THR manualized curriculum[16] is designed for beginner riders; (4) exceeding weight limits determined by the individual riding center sites' (Colorado and Maine) safety policies (5) vision, hearing, or physical impairments that would prevent participants from being able to be accurately evaluated (e.g. needing to see materials, see others, respond to questions) by study diagnostic measures or (6) unable to engage safely and independently with horses and others at the riding center. Of note, participants will not be allowed to begin baseline assessments until at least six months have passed from the time they last engaged in mounted EAAT, given pilot evidence for the six-month maintenance of THR effects. [112]

Study Design and Research Methods

(See Appendix for Calendar of Events/Timetable Document)

Setting: To address Aims 1 and 2, all THR and BA control interventions and physiological data collection will take place at two Premiere accredited PATH[17] international therapeutic riding centers: (1) Colorado: Hearts and Horses in Loveland and (2) Maine: Riding to the Top Therapeutic Riding Center in Windham. Note: study consent/assent and initial screening visits will occur at the following sites: Children's Hospital Colorado main campus and satellite clinic site locations, Colorado State University, and Maine Behavioral Healthcare/Center for Autism and Developmental Disorders clinic (CADD). The Colorado site is a large facility that has previously participated in our replication pilot[18] and is located in a rural area that draws referrals from northern Colorado and southern Wyoming where intervention resources are limited for youth with ASD who require extensive services to address psychiatric issues and related aberrant behaviors. The Colorado site will also recruit participants for the waitlist/hybrid groups to address exploratory Aim 3a & 3b. The Maine site is located in close proximity to a large inpatient and outpatient psychiatric treatment center for youth with ASD in Portland (Maine Behavioral Healthcare/Spring Harbor Hospital). Therefore, the Maine site will also recruit ASD youth psychiatrically hospitalized at Spring Harbor Hospital to address exploratory Aim 3c. Please note Maine site is closed as of December 31, 2023.

Recruitment: IRB-approved study recruitment materials will be used for participant recruitment in communities (i.e., schools, ASD parent-support networks, and service agencies) near each of the two therapeutic riding centers. The Maine Site will recruit participants discharged from Maine Health's Spring Harbor Hospital's inpatient and partial psychiatric unit. At the Colorado site, participants will also be recruited via the COMIRB-approved study flier sent through Children's Hospital Colorado MyChart EPIC messaging system.

Pre-screening: Pre-screening will take place via REDCap survey or phone interview. Pre-screen questions are the same regardless of modality. Pre-screen questions are specific to study inclusion and exclusion criteria previously described. Those meeting initial pre-screening inclusion criteria will be scheduled for an initial screening visit at their respective locations (Children's Hospital Colorado or Spring Harbor Hospital, ME).

Re-screening: Participants previously screened, consented and randomized, who have chosen to wait 6 months or longer to begin the intervention phase, will be re-screened with the ABC-irritability scale to ensure they still score ≥8 on that scale before they begin the intervention. Those not meeting ≥8 on the ABC-C -irritability scale will be excluded from the study.

Initial Screening (Visit I): As part of the informed consent/assent process, participants and their caregivers will be informed that they will receive modest monetary compensation for the pre-and post- speech therapy visits and for the six-month follow-up assessment visit (if randomized to the THR or BA group). They will also be informed that they will receive free 10-week interventions (THR, BA, or Hybrid). **The BA group will receive one free THR lesson following their completion of post-intervention assessments**. The youth's tolerance of study equipment (helmet, Shimmer3 EDA and ECG monitors with electrodes, and salivary cortisol collection swabs) follow successful procedures used in our pilot studies (e.g., social stories, modeling, desensitization). Participants will receive instruction on saliva-stimulating exercises using images and simulated chewing, as needed. Caregivers will be asked to complete the psychiatric diagnostic forms. Non-verbal IQ and ASD diagnostic screening assessments will be administered to youth to establish eligibility. If the youth meets screening criteria, the caregiver identified will be given the demographic forms to complete. The caregiver will be asked to identify a designated/consistent caregiver to complete pre-, during, post-and 6-mos assessment forms for reporter consistency. Participants meeting screening criteria will be randomized to study groups according to the randomization plan in the following section. Those participants will be given a riding center orientation packet to complete and return directly to the riding center prior to being scheduled for their riding center site screening visit (Visit II).

Randomization: For the Colorado site, the study statistician will generate random group assignment numbers stratified according to Leiter-III[106] nonverbal IQ standard score (>85 NVIQ and ≤ 85 NVIQ) to assign treatment (THR) or control groups (BA or waitlist/hybrid).After at least 24 participants have completed the waitlist/hybrid arm, this assignment will discontinue. Those assigned to the waitlist group will, after completing their 10-week waiting period and post assessments, participate in the Hybrid group (5 weeks BA and 5 weeks THR). For the Maine site, the study statistician will generate random group assignment numbers as specified above to assign to the treatment (THR) or control group (BA).

Riding center site screening (Visit II): THR and BA participants will complete a 45-minute -individual screening at their study site's riding center with a THR instructor. In the event that a THR or BA participant could not complete this riding center for behavioral reasons (e.g., feeling overwhelmed by the novel environment), participants will be allowed no more than two attempts for this screening. This will involve interviewing caregivers about participants' needs and abilities, caregivers completing the ABC-C Irritability screener, as well as engaging participants in a review of riding center safety rules, and practice with physiological equipment (i.e., putting saliva sampling rod in mouth and wearing the Shimmer3 EDA and ECG monitors) . For THR participants: Sit at an art table and engage in an art or puzzle activity about horses, enter the arena, engage in a 10-minute mounted activity on the horse, and then dismount. For BA participants: Sit at art table and engage in an art or puzzle activity about horses. Participants will then be assigned to a THR or BA group appropriate to their age and abilities. THR participants will be matched with a horse based on their unique needs consistent with standard practices of the THR industry.[17] Participants may be excluded from the study at this time due to exclusion criteria discussed earlier. The waitlist/Hybrid participants recruited for the Colorado site will follow both a combined version of BA and THR screening routine similar to the screening protocol previously described .

Time 0 (T0) Baseline intervention assessments (Visit III): Within one month before week 1 of the intervention phase, all participants (THR, BA, Waitlist, Hybrid) will be individually scheduled to complete the SALT[127] with the speech therapist at their study site (Colorado or Maine). The speech assessment will be audiotaped and transcribed to allow scoring of this measure. Video and audio files will be stored on a secure, password protected drive on the Colorado and Maine Health sites secure network servers. On the first day of the 10-week intervention (THR, BA, or Hybrid) or wait period, caregivers will complete the ABC-C,[98] SRS™-2,[126] EDI,[128] the WHOQOL-BREF [95] , and the Crisis Mental Health Care Usage Survey

If caregivers decide to postpone their child's 10-week intervention after they have already completed baseline intervention assessments (Visit III) and their child has completed only up to two intervention lessons, caregivers will be asked to complete an additional baseline assessment (caregiver forms only). The initial completed baseline assessments will not be used for data analyses. Additionally, when the participant is ready to complete their 10-week intervention, they will not attend the one or two lessons they may have previously attended so that they will only complete up to 10 intervention lessons as part of their participation in the study.

Riding center site 10-week intervention (Visits IV-XIII)*:* (**See Appendix for an overview of THR and BA interventions**).

Participants (THR, BA, Hybrid in Colorado only) will attend intervention sessions at the riding center in their recruitment state (Colorado or Maine). THR and BA control group times will be between 1:00-5:00 PM in order to collect salivary cortisol when levels typically decline.[111] Hybrid group times will also occur within this same afternoon time period for consistency, even though salivary cortisol samples will not be collected. Intervention groups (THR, BA, Hybrid) will consist of 2-4 participants, be led by a THR instructor, follow a consistent routine, and learn horsemanship skills via activities tailored to ASD learning styles outlined in the THR manual.[16] Each week of the intervention phase (THR, BA or Hybrid), caregivers will complete the ABC-C,[98] rating their child's behaviors over the past week prior to the current lesson. THR Intervention: Each participant will have an assigned horse and volunteer(s) (one horse leader and up to two side walkers). Similar to the RCT[13], every effort will be made to enable each participant to ride a consistent horse throughout the study. Barn Activity Intervention: Each participant will have one assigned volunteer and will have no contact with horses at the riding center, just view horses at a distance. There will be a life-sized stuffed horse in the BA group for hands-on learning related to the weekly topic.

For THR and BA participants only: To increase compliance with mediator assessments, each week prior to starting the groups, participants will follow a consistent routine of wearing both their Shimmer3 EDA and ECG devices. The Shimmer EDA device will be worn on the wrist of participant's non-dominant hand and the Shimmer3 ECG cardiac monitor will be adhered to the chest wall area using adhesive electrodes. Participants will sit at a group art table with their respective small group (THR or BA) while study personnel will instruct participants to place the 10cm long foam swab rod under their tongue for one minute while watching a 1-minute timer. THR participants will then don their riding helmets and enter the riding arena and the BA will begin their group activity at the group table. Each week after conclusion of the THR and BA interventions, participants will sit with their respective groups (THR or BA) at an art table for 5 minutes followed by another saliva sample.

Time 1 (T1) Mediator assessments taken weekly during the 10-week intervention, but will only be analyzed for weeks 1, midpoint, and 10. Physiological data (HR, HRV, and EDA) measurements will be taken continuously during the entire intervention lessons. Sections of analysis will be segmented from: (a) Three to five-minute seated art activity baseline before either intervention; (b) 45-minute THR riding or BA activity; (c) 15-minute THR grooming and tacking or BA scrapbook activity; and (d) 5-minutes post intervention that includes a seated art activity. RAs will annotate time and relevant activity levels for each period to correspond to device timestamps. Salivary cortisol samples will be taken during the five-minute seated baseline before THR or BA intervention and again 5 minutes post intervention. Caregivers will complete the Physiological Sample Collection Log each week of the 10-week intervention.

Time 2 (T2) Post- intervention assessments (Visit XIV)*:* On the last day of the 10-week intervention (THR, BA, Hybrid) or waitlist period, caregivers will complete the ABC-C,[98] SRS™-2,[126] EDI,[128] , the WHOQOL-BREF [95] , and the Crisis Mental Health Care Usage Survey. Within one month of the conclusion of week 10 of the intervention phase, all participants (THR, BA, waitlist, and Hybrid) will be individually scheduled to complete the SALT[127] with the same speech therapist they saw for the Pre-intervention SALT at the Colorado or Maine hospital site. If there is a speech therapist personnel change from a participant's pre- to post- intervention SALT assessment visit, this change will be noted in the REDCap database to control for personal novelty bias in the analysis.

Time 3 (T3) Six-month post-intervention assessment (Visit XV)*:* Only THR and BA participants will be individually scheduled to complete SALT[127] with the speech therapist at their study site (Colorado or Maine). During this visit, caregivers will complete the ABC-C,[98] SRS™-2,[126] EDI,[68] , the WHOQOL-BREF [95], and the Crisis Mental Health Care Usage Survey. If there is a speech therapist personnel change from a participant's post- to 6-month post- intervention SALT assessment visit, this change will be noted in the REDCap database to control for personnel novelty bias in the analysis.

Post study survey with study intervention staff: The Intervention Team Reflection Survey will be sent to the study coordinator at both riding centers (CO and ME) with a request to send out to instructors and volunteers who participated in the study.

Description, Risks and Justification of Procedures and Data Collection Tools:

Informed Consent: If all protocol initial phone-screening criteria are met, the youth and their parent/caregiver/guardian will be scheduled for an initial study screening visit at their respective recruitment site at CU Anschutz./Children's Hospital Colorado, CSU, or Maine Behavioral Healthcare. At that screening visit, the study personnel will review consent verbally, highlighting important aspects of the consent with parent/guardian(s); allow time for them to read the consent thoroughly; review/explain any aspects of the consent form or study that the parent/guardian may have questions about. If the parent/guardian has questions about the study that the study personnel is unable to

answer, the study personnel will ask the PI, Dr. Gabriels, or subcontract site PI, Dr. Siegel, at their respective recruitment site (CU/CHCO, CSU or ME) to speak with the parent/guardian. It is anticipated that many of the participants recruited for this study will have a mental age of less than seven years, and this will be determined by the PI or subcontract PI in conjunction with the caregiver. Given the nature of social and communication difficulties specific to ASD (e.g., understanding the social nuances of language), just having a chronological age of 7 years does not necessarily mean that the child/adolescent will understand the assent. Therefore, the best and most conservative way to ensure that the child/adolescent is able to understand the assent is to require that all potential participants over the age of 7 years, who would be required to sign the assent, are verbal and have an IQ in the low average range or above (i.e., IQ score $\geq$ 80). In the event that a child/adolescent does have a mental age equal to or above the age of 7 years, then the assent form will be explained to the child/adolescent and they will then be asked to sign this form. If the child/adolescent is able to understand the assent document and chooses not to provide their assent, the parent or legal representative's wishes will not override those of the child/adolescent, and they will not participate in the study. A copy of the consent form (and assent form, if applicable) will be provided to the caregiver.

Postcard Consent: Study intervention personnel will be consented via a postcard consent which explains why they are being asked to participate in the Intervention Team Reflection Survey. The consent and the survey will be sent to the study site coordinator to distribute to all site personnel involved in the study interventions. On this form, intervention personnel are instructed to acknowledge their consent to use their direct quotes anonymously in future study publications. Study intervention personnel are instructed to avoid using any personal identifying information on the survey. Finally, intervention personnel are instructed to return this form to the study site coordinator who will return to the overall study PI, Robin Gabriels PsyD.

Duration of Participation:

Participants will be asked to participate in this study for up to 11 months, depending on their group randomization. During this time and depending on their group randomization status, participants will attend up to 21 study visits described as follows:

- Two initial consent/screen visits (diagnostic and cognitive assessments, intervention riding center site screening)
- One pre-intervention evaluation at the Colorado or Maine subcontract site (within one month) prior to the 10-week intervention (THR/BA or Hybrid) or waiting period. *Note, **if caregivers decide to postpone their child's 10-week intervention after they have already completed baseline intervention assessments (Visit III)** and their child has completed only up to two intervention lessons, caregivers will be asked to again complete the baseline assessment (caregiver forms only). The initial completed baseline assessments will not be used for data analyses.*
- Ten intervention visits at the riding center (THR, BA or Hybrid group lessons)
- One post-intervention evaluation at the Colorado or Maine subcontract site (within one-month post intervention for THR, BA, Waitlist and Hybrid group participants). Note: The post-waiting period evaluation of the Waitlist group may also serve as the pre-intervention evaluation for their involvement in the Hybrid group unless there is a two-month lag period before the start of their Hybrid group. In such case, the Hybrid group will complete an additional pre-intervention evaluation within one month of the 10-week Hybrid intervention.
- One post-intervention evaluation at the Colorado or Maine subcontract site six-months post intervention for THR or BA group participants only

Recruitment and Retention Plan:

For this proposed study, we will localize recruitment efforts in communities (e.g., hospital psychiatric treatment centers, schools, respite care, local parent autism societies, and mental health care agencies) in the surrounding areas closest to the riding centers in northern and southern Maine regions. Please see Facilities and Resources document for a table to referral sources in those communities.

Retention plan for study assessment periods:

At the Colorado site, we plan to offer an option to participants to conduct the initial screening visit and the initial pre- post- and six-month post-evaluation visits at either a Children's Hospital Colorado site Maine Behavioral Healthcare. At the Maine site, the diagnostic/screening visit, pre- post- and six-month post-evaluation visits will generally be conducted at the Center for Autism and Developmental Disorders, Westbrook ME. If necessary, to accommodate a family, screening and assessment visits may be conducted at the riding center in Windham, Maine.

- For the current proposal, we anticipate lower attrition than the 20% rate in this pilot for the following reasons:
  - Outcome evaluations will occur in closer proximity to participants' residence
  - THR, BA (both sites), and Hybrid (CO only) groups will receive FREE 10-week riding and/or barn activity group lessons.

- The BA control group will receive one free riding lesson after their 10-week intervention period.
- Increased monetary incentive will be provided for participants in the THR and BA groups for the 6-month follow-up.

Participant Incentive Costs:

- Baseline and post intervention (THR, BA, or Hybrid) or waiting period speech therapy visit= $20.00 x 2 visits.
- Maine Site Only: Each intervention day the participant is in attendance to help support gas costs = $10.00 x 10 visits. We will disperse payments at Week 5 and Week 10 in accordance with the participant's attendance for weeks 1-5 and weeks 6-10 respectively.
- Colorado site: Monetary support for gas costs to the riding center for intervention weeks 1-10 may be available to participants, but eligibility will be determined by study personnel.
- Six-month follow-up assessments for THR and BA groups only = $30.00 x 1 visit
- Note: All participants will receive FREE interventions (THR, BA, and Wait-list/Hybrid groups). The BA group will receive one FREE riding lesson after they complete their 10-week intervention assessments.

Therapeutic Horseback Riding at both study sites (CO and ME): Participants will be required to have a medical clearance form completed and signed by their primary care physician prior to engaging in the THR, BA, or Hybrid activities at the riding center. Horseback riding has inherent risks and this study involves weekly horseback riding lessons. Risks from horseback riding could include fall, loss of consciousness, bodily injury, dismemberment or death. These risks are mitigated by this study's use of a PATH International Premiere Accredited Center. This means that they adhere to national standards to ensure safe and effective operations throughout all areas of programming. More information about these center accreditation standards can be accessed by visiting the PATH International website at www.pathintl.org. In addition, PATH international certified instructors will be conducting the therapy. These instructors have demonstrated PATH competency in the areas of equine management, horsemanship, riding instruction, teaching methodology and knowledge of disabilities. These instructors also have Cardiopulmonary resuscitation (*CPR*) and First Aid certifications. **(See riding centers' participant rider safety standards and procedures in the Appendix documents.)**

Assessments: This study also involves caregivers completing questionnaires and participants being assessed with diagnostic and cognitive measures. Any results from these assessments of the child, conducted as part of this study, may be uncomfortable for the caregivers to hear; however, these results will be discussed with caregivers in detail, and a written report of the cognitive assessment results provided upon request of the caregiver. The study assessments take a considerable amount of time; therefore, the NVIQ and ADOS-2 assessments may be divided into sessions on different days in order to reduce fatigue and increase compliance. The PI and other study evaluators have extensive experience in the administration of assessments with the proposed study population.

Treatment Randomization and Termination: Participants at the Colorado site will be randomly assigned to either a THR group, BA control group or waitlist group. Participants at the Maine subcontract site will be randomly assigned to either a THR group or BA control group. BA control participants who are not randomized in the THR group may experience frustration if they are not able to interact with horses as part of their involvement in the study. However, participants in the BA control group will be offered the opportunity to receive a one-hour free horseback riding group lesson at their respective sites' riding center that will occur after they complete 10-week post-intervention SALT assessment. Other than the BA activity 1 free THR lesson, BA and THR participants will need to refrain from riding horses until their 6-month post-assessment is completed. Participants randomized to the 10-week waitlist (CO site only) will then be involved in the Hybrid group that included five weeks of BA followed by five weeks of THR. However, the waitlist participants may experience disappointment that they are not immediately able to interact with horses. The THR, BA, and Hybrid groups may view the termination of the study period as a loss of treatment benefits. If participants wish to continue THR beyond the study period, the therapeutic riding sites (CO or ME) have availability for those who wish to continue these interventions on a fee for service basis.

Violation of Privacy and Loss of Confidentiality at both study sites (CO and ME): These are both risks to which research participants are exposed. The possibility of these risks increases when protected health information is collected. The protected health information that will be collected include: the participant's name, date of birth, demographic information (age and sex), diagnosis, medical history and treatment information, psychological testing results, and survey questionnaires regarding the child's behavior. This information will be available to study staff only for the purposes of data analysis and evaluation of this research project. Study staff will have completed certifications (CITI Basic Course and HIPAA Research Course) prior to working on this study. All participants' data will be coded with a number and the key for this coding will be kept in a separate file from the data file and will be password protected. Personal identifiable information will be removed from data stored electronically and access will be limited to personnel associated with the project. Participants' age will be derived by subtracting their DOB from the date the caregiver completed the assessment

tools but will not be included as part of the participant data file. All assessment tools will be de-identified (i.e., participants' names and DOB will be deleted from these forms). Participants' files will be coded with a number and kept in a locked file cabinet. The speech assessment will be audiotaped and transcribed to allow scoring of this measure. Video and audio files will be stored on a secure, password protected drive on the Colorado and Maine Health site's secure network servers.

Reactions to Environmental Allergens: Some participants may be allergic to pollens, animals or other seasonal or environmental factors that are present in the outdoor environment of the riding centers (CO and ME). Caregivers will be asked about environmental allergies that the participant may have and what sort of allergic reaction the participant typically displays. Caregivers will be encouraged to monitor their child for any signs of allergic reaction and to bring with the participant any necessary medications to decrease their allergic reaction, should such develop. Caregivers will be encouraged to take their child to be evaluated by their Primary Care Provider, if allergy symptoms persist.

Adequacy of Protection Against Risks: The risks to participants regarding THR are mitigated by this study's use of two sites that are PATH Intl.[17] Premiere Accredited Centers. Première accreditation status means they follow rigorous standards and comply with site visits every five years to assess the sites' infrastructure as well as ensure safe and effective operations for horses and humans throughout all areas of programming. *More information about these center accreditation standards can be accessed by visiting the PATH Intl. website at www.pathintl.org.* The standard procedures used by the therapeutic riding centers for this proposal decreases risks to the riders and staff implementing safety policies standard in the field and according to PATH Intl. guidelines. Such safety standards include the following: requiring participants wear protective gear, monitoring the temperament of the horses, training horses to participate in therapeutic horseback riding experience with children who have a variety of disabilities, and instructing participants about safe practices around horses. In addition, PATH Intl. certified instructors will be conducting the therapy. These instructors have demonstrated PATH Intl. competency in the areas of equine management, horsemanship, riding instruction, teaching methodology and knowledge of disabilities. These instructors also have Cardiopulmonary resuscitation (*CPR*) and First Aid certifications. **(See riding centers' participant rider safety standards and procedures information in the Appendix documents.)**

Participant Discontinuation Criteria: Participants who miss more than two out of the 10 THR, BA or Hybrid intervention lessons will automatically be dropped from the study. Participants who develop significant allergic reactions to conditions at either of the riding centers (CO or ME) will, at the discretion of the caregiver or the riding center staff, be dropped from the study. There are no protocol-stopping criteria as there are no *a priori* hidden risk concerns from which to build criteria. The participant discontinuation criteria will adequately protect study participants.

Data and Safety Monitoring Plan: The Principal Investigator (PI) at the Colorado site will monitor data and safety and be responsible for promptly reporting any related protocol concerns, including Serious Adverse Events (SAEs) and Adverse Events (AEs), to the Research Participant Advocate and to COMIRB. The Maine site and Maine Health IRB will rely on the COMIRB for review, approval, and ongoing regulatory oversight of the study. The PI at the Maine site will monitor study conduct at the site, monitor data and safety, and will be responsible for promptly reporting concerns, SAE and AE to the Colorado lead PI and to COMIRB as the reviewing IRB of record. The Principal Investigator at the Colorado and Maine subcontract site have no conflicts of interest (financial, academic, professional) that might interfere with being responsible for data and safety monitoring. The data will be examined in the course of the protocol on at least a monthly basis by the Principal Investigator at the Colorado and Maine subcontract site. Access to the database will be restricted to study personnel and standard conventions to protect the security of the database will be observed. Additionally, the PI (Colorado site) and subcontract PI (ME) will report any SAEs for this study that might include a child falling from the horse or the possibility of injury to one of the study personnel during the evaluations or treatments. The Principal Investigator (CO) and subcontract PI (ME) will also routinely communicate with the study personnel at their respective sites (Colorado and Maine) regarding the occurrence of any SAEs and be responsible for immediately communicating this information to their respective IRB Research Subject Advocate.

Study data will be collected and managed using REDCap (Research Electronic Data Capture).[131] REDCap is a secure web application designed to support data capture for research studies, providing user-friendly web-based case report forms, real-time data entry validation (e.g. for data types and range checks), audit trails and a de-identified data export mechanism to common statistical packages (SPSS, SAS, Stata, R/S-Plus). The REDCap system was developed by a multi-institutional consortium which includes University of Colorado–Denver and was initiated at Vanderbilt University. The database is hosted at the University of Colorado–Denver Development and Informatics Service Center (DISC), which will be used as a central location for data processing and management. REDCap data collection projects rely on a thorough study-specific data dictionary defined in an iterative self-documenting process by all members of the

research team with planning assistance from the *DISC*. This iterative development and testing process results in a well-planned data collection strategy for individual studies. REDCap also includes a powerful tool for building and managing online surveys. The research team can create and design surveys in a web browser and engage potential respondents using a variety of notification methods. REDCap is flexible enough to be used for a variety of types of research and provides an intuitive user interface for database and survey design and data entry.

The study database for this multisite study will be housed at CU Denver, and access to the study database will be monitored and controlled by the study PI at the Colorado site. All study users will be required to complete applicable training to gain access and will each have a unique logon and password. The data will be backed up daily. There will be an audit trail of users and changes to the data.Maine site's, physiological data (EDA, HR, HRV) will be stored and managed through OpenText, a secure, online, data sharing platform.

Data Analysis Plan:

***Data Sharing Plan**:* Study data will be collected and managed using REDCap electronic data capture tools hosted at the University of Colorado.[131] REDCap (Research Electronic Data Capture) is a secure, web-based application designed to support data capture for research studies, providing: (1) an intuitive interface for validated data entry; (2) audit trails for tracking data manipulation and export procedures; (3) automated export procedures for seamless data downloads to common statistical packages; and (4) procedures for importing data from external sources

Physiological data (HR, HRV, EDA) will be monitored on an ongoing basis for quality control throughout the study by physiological activity assessment experts , both visually and using automated quality control software.[132] Video recordings of 20% of each 10-week intervention will be shared with the intervention fidelity coordinators to rate intervention fidelity. Colorado's site physiological and fidelity video data will be shared through the CU One Drive and Maine site's physiological and fidelity video data will be shared through OpenText. Fidelity videos may also be shared via flash drive mailed via FedEx.

Data auditing and cleaning: HRV: Following the authoritative guidelines given in the Task Force,[133] data will be band-passed filtered (0.5-45Hz), heartbeats will be detected by automatic multiscale-based peak detection,[134] and lost data and/or premature atrial/ventricular contractions corrected via cubic spline interpolation. Time-domain indices (HR; SD; root-mean square of successive differences, RMSSD) will be calculated per standard methods. Frequency-domain indices (LF, low-frequency power; HF, high-frequency power) will be calculated via Lomb-Scargle Periodogram, and reported as natural log values. Standard covariates will be applied, and data quality reporting indices will be reported.[135,136] EDA will be decomposed into tonic and phasic activity using standard methods [40] and de-noised depending on the accelerometer data – periods of high physical activity will be removed from the analysis period. The mean, SD, and non-specific skin conductance responses (NSSCRs; bursts of activity over 0.2uS) will be extracted per time period. NSSCRs will be reported as depolarizations per minute. Signal quality for both physiological records, which is frequently problematic during ambulatory recording, will be maintained through the combined use of automated tools for quality control,[132] visual artifact identification, and redaction.[137]

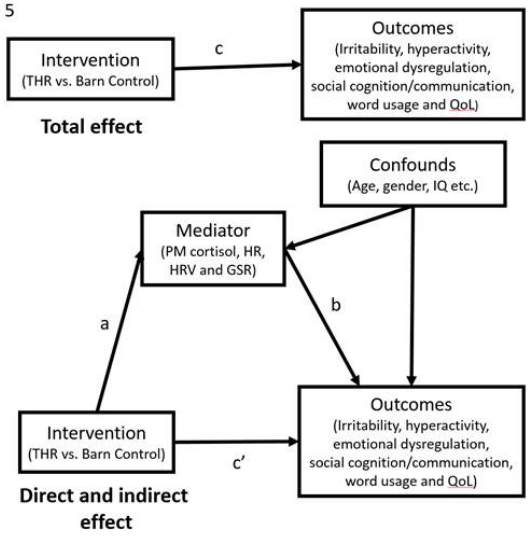
During each 10-week intervention (THR and BA), Dr. Cory Smith's 's lab team will meet with study PI, Dr. Gabriels and RAs from Colorado and Maine sites to monitor the Shimmer ECG (HR/HRV) and EDA data collected for quality and for possible equipment failure issues. Dr. Cory Smith's lab team in collaboration with Dr. Gabriels will then determine which data points collected will be used for analyses, ensuring that each subject has a beginning point (Lesson 1 or alternate lesson 2 or 3), midpoint (**Lesson 6** or alternate lesson 7), and endpoint = (**Lesson 10** or alternate lesson 9 or 8).

General consideration of statistical analysis: Prior to the start of any formal analyses, data for each variable will be examined to identify unusual values that need to be queried, including outliers, patterns of missing values, and non-Gaussian distributions (i.e., have significant skewness and kurtosis). Baseline demographic and clinical characteristics will be summarized using descriptive statistics and grouped by intervention arms. Between-arm difference in demography and baseline variables will be examined using t-test or chi square test respectively for continuous and categorical variables. Any imbalance that could potentially be a confounding factor for the outcome will be adjusted for in statistical models (see below). The primary efficacy analysis will use the "intent-to-treat (ITT)" principle to test our study hypotheses under practical and realistic conditions where not all participants follow or complete the program. All randomized participants will be analyzed in the ITT analyses. Linear Mixed Effect Model (LMM) will serve as the primary method to handle missing values. Other analyses such as completer analysis and analyses with imputed data will be deemed as sensitivity analysis for efficacy examination. However, for mediation analysis, only participants with both the

mediator (intervening) variable and outcome variables available will be included to explore mechanism of the intervention effect. No interim analysis will be conducted. In all analyses, a $p \leq 0.05$ will be considered significant. SAS and Mplus software will be used for these analyses.

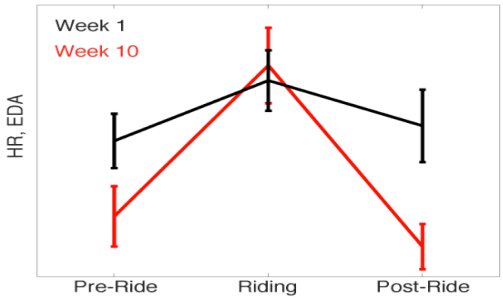
AIM1**: **Efficacy assessment (for H1a-1c)*:* LMM with unstructured covariance will be used to model baseline and end of treatment (EoT or T_2) primary and secondary outcome measures as well as expanded outcomes where fixed effects include evaluation time (baseline or EoT) and intervention arm indicator (THR/BA), as well as their interaction. Estimate of the interaction will be used to quantify efficacy (i.e., effect size) and tested for statistical significance. **Mediation analysis** (H1d): **Fig. 5** illustrates our mediation model. For ease of interpretation, the change from T_0 to T_2 will serve as the outcome and change from T_0 to T_1 the outcome will serve as the potential mediator variable. Consistent with the previous trial recommendations for the analysis of mediation in RCT studies, we will use product of coefficients tests for the intervening variable effect[136] to examine statistical significance of mediators on a completer analysis basis. Specific procedures include: [Step 1] Conduct a simple regression analysis with the intervention predicting the outcome (Path c). This will reconfirm the THR efficacy from previous trial using the LMM discussed above [Step 2]. Among the outcome variables significant on path c, conduct a simple regression analysis with intervention predicting the potential mediator to test the path a. [Step 3]. Among the significant potential mediators in path a, conduct multiple regression analysis with the intervention and the mediator predicting outcome.

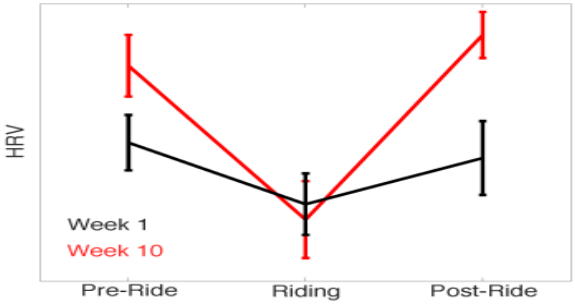
Figure 5



Based on the last two models, the Indirect effect (i.e., intervention→mediator→outcome) can be calculated as the product of the coefficients of paths a and b and tested using the Bootstrap method. Significance of indirect effect would support that the THR effect on the outcome is, to a certain extent, mediated through the mediator variable. We initially will conduct this modeling separately by mediator and outcome. We will build a multiple path model among significant mediators found from these separate analyses. Because randomization of treatment does not necessarily control for potential confounding variables on the mediator effect on the outcome (path b), we will also perform mediation analysis with adjustment for age, gender, IQ, and baseline to add robustness to our analyses.

Finally, see Figs. 6 and 7 for expected results of physiological mediation variables as hypothesized in Aim1c: *H1c. THR group will have a decrease in heart rate (HR) and electrodermal activity (EDA) and an increase in heart rate variability (HRV), both indicating a calm state, from baseline to week 10 of the intervention."*





***AIM 2**:* LMM model will be used to assess the efficacy of THR on the outcomes found significant from efficacy analysis for Aim 1. This LMM model consists of time [baseline, end of intervention (EoT), six-month measures (T_3)], and intervention arm (THR or BA), as well as their interaction term as fixed effects and an unstructured covariance. Between-arm and within-arm multiple comparisons will be conducted should there be significant overall time by treatment interaction. Between-group difference in the change from baseline to T_3 will be used to assess the long-term efficacy of THR on T3 outcome. The primary analysis will be ITT and no data imputation will be used. However, we will conduct two sensitivity analyses to examine robustness of primary analysis results. First, we will use the Markov Chain Monte Carlo (**MCMC**) method to impute any missing values at T_2 or T_3 and then fit the same LMM model. Second, the same LMM model will be applied to completers only.

***Exploratory AIM 3**:* To assess whether BA affect outcomes (H3a), we will first analyze the 60 Colorado participants randomized with 1:1:1 ratio into waiting list (followed by Hybrid intervention), BA or THR group. Next, we will compare the 20 wait list participants to all participants in the BA and THR groups respectively. To assess the effect of the

hybrid intervention (H3b), data during the hybrid treatment period will be compared to within-person outcome data during the wait period. Moreover, we will compare hybrid treatment participants to all BA and THR participants. For both hypotheses 3a and b, LMM models similar to the one for Aim 1 efficacy analysis will be used. Again, the between-group difference in outcome change between baseline and 10 weeks will be used to assess efficacy and tested for statistical significance. The 95% confidence interval for efficacy and corresponding effect size (Cohen's D) will be reported. Any imbalanced confounding variables will be adjusted in the model. To explore whether THR effects are any different between participants recently discharged from the psychiatric hospital and those without recent psychiatric hospitalization, we will compare THR effects between these two groups of THR/BA participants by introducing three extra terms in the model for Aim 1. They include an indicator variable (RecentHosp) for this sub-classification (with or without recent hospitalization) and the intervention*RecentHosp and intervention*RecentHosp*time interaction terms. Contrasts will be set up to test the effect of interest and 95% CI for this effect assessment will be calculated

***Power analysis**:* Because significant intervention effect is the prerequisite of mediation analysis, this study is powered primarily on analyses of the primary outcome, ABC-C irritability. For the primary efficacy analysis at T2, effect sizes were estimated for the primary outcome (ABC-C irritability) based on data from our THR RCT[13] among participants who had psychiatric diagnosis (N=27 for THR and N=26 for BA, data not published). The irritability score decreased by 6.9 (1.66 SEM) in the THR group and 2.3 (1.8 SEM) in the BA group, resulting in a Cohen's d of 0.62 for efficacy. We will enroll 142 participants (71 per arm). To allow for 20% attrition (which is 1% higher than that in our previous trial) at T_2, the sample size for completer or mediation analysis would be 56 per arm. In this case, we will have 97% power at 5% significance to detect significance of THR efficacy for ITT analysis (Aim1) and 90% power for completer analysis. For the mediation analysis, our previous data indicate that the standardized coefficient for the total effect (path c) is 0.34. There are no previous data available for estimating coefficients for paths a and b. We assume that both coefficients are of medium effect size[137] and the indirect path will explain half the total effect. That is, a×b=0.17 (e.g., a=0.4, b=0.43). In this case, a total sample size of 112 provides 90% power at 5% significance to detect a significant indirect effect. This power estimate is based on the Sobel test and was calculated by the R package, power Mediation. **Table 2** shows the power for analyzing secondary and exploratory outcomes at T_2 if the ES is similar to those from data in our previous RCT. (Based on 18 THR and 11 BA, participants with concomitant psychiatric diagnosis and six-months follow-up ABC-C data from our RCT, the ES for THR effect on the six-month irritability score was 0.45 (not published) as compared to BA participants. In this case, we will have 81% power for the primary ITT analysis (n=81 per arm) to detect the THR effect on ABC-C, but 56% power to see statistical significance for completer analysis. This analysis assumes that 45 participants per arm will be followed six-months after the intervention. Aim 3 is more hypothesis generating and exploratory. 20 extra participants will be enrolled in the wait list group who subsequently received Hybrid intervention. To minimize the imbalance between groups, we will randomize with 1:1:1 ratio the first 60 Colorado participants into BA/THR/wait list group. The minimum detectable effect size needs to be 0.9 (0.74) Cohen's D to have 80% power if 20 wait list (or Hybrid) participants will be compared to 20 (or 56) other participants.

TABLE 2	Power at 5% significance	
Power from ES	ITT	Completers
	(n=81/arm)	(n=56/arm)
ABC-C Irritability	***97***	***90***
ABC-C Hyperactivity	94	83
SRS Social Cogn.	92	80
SRS Social Comm.	>99	>99
SALT # diff. words used	>99	>99
SALT # words used	>99	>99

Supplementary analyses: 1.) It is likely that participants will have changed psychoactive medication during the 10- week intervention period and this change is not balanced between two study arms. History of medication change for each participant will then be documented and characterized into few categories, as appropriate. This category of medication change (to be determined) will be introduced into the statistical models for Aims 1 and 2 to adjust for their potential confounding effect for efficacy and mediation analysis. 2.) to examine the effect of intervention sites (Colorado or Maine), we will tabulate the effect size of THR effect by site and introduce site and site by intervention terms into the models mentioned above for efficacy analysis in order to statistically test the site by intervention interaction effect. 3.) We will explore factors that related to the response to THR intervention traditional variable-centered and the state-of-the-art person-centered methods. We will classify each THR subject as responder or non-response (with criteria to be determined) and use logistic regression to identify baseline demography factors and clinical factors (such as psychological diagnosis) that are associated with the response. Moreover, latent profile analysis will be used to group all the participants into few clusters who share similar demography and baseline clinical characteristics. Difference in intervention effect will be examined across these clusters to determine who are more responsive to the intervention. In the later analysis, the outcome will be change score after intervention, the model will consist of cluster, treatment and their interaction. Test of interaction is of primary interest. 4.) In our previous R01 THR study, we observed that the intervention (Barn or THR) reached it.

Summarize Knowledge to be Gained:

Our proposal aims to moves the HAI field beyond efficacy trials, specifically regarding THR:

The experiments outlined in this application are designed to move the HAI field forward to provide evidence to begin to answer the more nuanced questions of **how** it works, **for whom** does it work and for **how long,** as well as **under what conditions** (dose) does it work. The historical theoretical focus on HAI leading to arousal regulation/conditioning is an ideal target for mediation analyses via physiological measurement. The knowledge to be gained from this proposal will provide evidence for physiological response patterns as a causal pathway to THR outcomes and has the potential to guide future researchers who wish to explore mechanisms by which HAI demonstrates benefits beyond THR and to other populations.

Aim 1 addresses the call to identify "causal relationships" and evaluate the "community-level and social benefits" of HAI, as identified by the PAR 18-213 HAI funding announcement. The proposed ASD study population tends to have atypical arousal response profiles, particularly youth with ASD and psychiatric diagnoses, making them more likely to benefit from an intervention such as THR whose outcomes may be mediated by physiological arousal regulation benefits. In line with the NIH RDoC framework,[93] elucidating physiological pathways influenced by THR in ASD may also suggest how THR could be useful for other psychiatric populations where similar psychophysiological pathways are known to be impaired. We also propose to expand our outcome measures to include evaluation of wider ranging benefits of THR, as children with ASD are at high-risk for behavioral disturbances that can impair their QOL as well as their caregivers. National surveys and studies indicate higher rates of caregiver stress and frustration related to the care and management of behaviors of children with ASD compared to other special needs populations,[7,8] suggesting this is a critical outcome to measure multidimensionality, including caregivers' physical, mental, and social well-being.[94] We propose to use the *World Health Organization's Quality of Life Instrument (WHOQOL-BREF)* [95], which is used for a variety of populations, including ASD, to provide an index to measure caregivers' adjustment.[94] Our inclusion of this standardized QOL measure will allow us to broaden the understanding of HAI efficacy to other research disciplines as well as understand the impact of THR on caregivers QOL. Problems with emotional dysregulation are present in other psychiatric populations (e.g., anxiety and mood disorders) and not specific to the ASD population.[37] Our inclusion of a distinct outcome measure of ER that is sensitive to change, quantifies variability in emotion dysregulation, has been validated for use with an ASD population with a variety of ability levels, and has the potential to generate information needed for follow-up studies broadens examination of THR as a potentially effective transdiagnostic intervention.

Aim 2 determines if THR outcomes are maintained over time. We propose to address the paucity of HAI research examining long-term maintenance of effects. In addition to our six-month follow-up report[18] in a subset of participants in our THR RCT[13], to our knowledge, only one other study to date has attempted to prospectively examine residual effects of THR in children with ASD, but it had several methodological limitations including teacher-report measures, lack of a control condition, and an unplanned six-week break during the initial treatment phase.[96] Information on longer term THR effects has important clinical practice implications, including the possibility that THR could be an adjunctive intervention to the current practice of using psychotropic medications to reduce symptoms of aberrant behaviors in ASD youth.[74]

Aim 3 explores dose and sub-population effects of THR and Barn Activity (BA) interventions. This aim addresses important next steps to advance the HAI field of research. A recent systematic review of HAI research with the ASD population discussed the field's tendency to employ diverse research methods that lack replication and use of non-homogenous ASD samples, which limits the ability to determine who might benefit the most from HAI and under which conditions.[71] Exploratory analyses of our RCT[13] sample showed stronger positive outcomes in the subgroup of THR participants with co-occurring psychiatric disorders, which provides support to focus on this group who stands to benefit the most from THR. Additionally, the BA no-horse control group in our RCT[13] showed significant within-group improvements in irritability and hyperactivity behaviors at the conclusion of the intervention, but we were unable to draw conclusions about the effects of that BA group due to the absence of a nonintervention control. Thus, we aim to both explore the effects of BA compared to a waitlist control and explore the dosing effects of a Hybrid group (i.e., 5 weeks BA followed by 5 weeks THR).

Our project design is an innovative approach that will provide ground-breaking advancements in the HAI field (specifically THR). Our approach innovates by: conducting a large RCT targeting a majority subset of the ASD population, who have indications of being potentially stronger THR responders[99]; using an intervention manual for both experimental and control conditions and measuring their fidelity to ensure implementation uniformity, specifically accounting for effects of the horse while controlling for social attention throughout the study; including a method to understand the six-month maintenance of THR outcomes vs. a control; and implementing the study in two community-based rural settings to expand the ecological validity of our manual-based THR intervention.[16]

We are grounding our mediation analysis in the framework of a model to be tested using a RCT. Methodologically, including multiple physiological measures will allow us to decrease the biasing effects of measurement error[58] and more systematically assess whether positive THR outcomes are partially accounted for by physiological response profiles over time in youth with ASD and co-occurring psychiatric diagnoses.

Our measurement approach is technically innovative. Our proposal seeks to employ objective measures of participants' experiences by using unobtrusive state-of-the art, ambulatory physiological measures during an HAI activity to capture the potentially complex profile of mediating factors that may explain previous outcomes. Our methodology takes advantage of technological advances in wearable physiological biosensors that overcome the inherent difficulty obtaining reliable self-reports on emotional and physiological states in individuals with ASD, a portion of whom are either minimally-verbal or have alexithymia.[103-105]

References:

1. Zablotsky B, Black LI, Maenner MJ, Schieve LA, Blumberg SJ. Estimated Prevalence of Autism and Other Developmental Disabilities Following Questionnaire Changes in the 2014 National Health Interview Survey. *National health statistics reports.* 2015(87):1-20.
2. Amendah D, Grosse SD, Peacock G, Mandell DS. The Economic Costs of Autism: A Review. In: Amaral DG, Dawson G, Geschwind DH, eds. *Autism Spectrum Disorders.* New York: Oxford University Press; 2011:1347-1360.
3. Salazar F, Baird G, Chandler S, et al. Co-occurring Psychiatric Disorders in Preschool and Elementary School-Aged Children with Autism Spectrum Disorder. *Journal of autism and developmental disorders.* 2015;45(8):2283-2294.
4. Simonoff E, Pickles A, Chairman T, Chandler S, Loucas T, Baird G. Psychiatric disorders in children with autism spectrum disorders: Prevalence, comorbidity and associated factors in a population-derived sample. *J Am Acad Child Adolesc Psychiatry.* 2008;47(8):921 - 929.
5. Jahromi LB, Bryce CI, Swanson J. The importance of self-regulation for the school and peer engagement of children with high-functioning autism. *Research in Autism Spectrum Disorders.* 2013;7(2):235-246.
6. Cadman T, Eklund H, Howley D, et al. Caregiver burden as people with autism spectrum disorder and attention-deficit/hyperactivity disorder transition into adolescence and adulthood in the United Kingdom. *J Am Acad Child Adolesc Psychiatry.* 2012;51(9):879-888.
7. Lecavalier L, Leone S, Wiltz J. The impact of behaviour problems on caregiver stress in young people with autism spectrum disorders. *J Intellect Disabil Res.* 2006;50(Pt 3):172-183.
8. Schieve LA, Blumberg SJ, Rice C, Visser SN, Boyle C. The relationship between autism and parenting stress. *Pediatrics.* 2007;119 Suppl 1:S114-121.
9. Mazefsky CA, Borue X, Day TN, Minshew NJ. Emotion regulation patterns in adolescents with high-functioning autism spectrum disorder: comparison to typically developing adolescents and association with psychiatric symptoms. *Autism Res.* 2014;7(3):344-354.
10. Beetz A, Uvnas-Moberg K, Julius H, Kotrschal K. Psychosocial and psychophysiological effects of human-animal interactions: the possible role of oxytocin. *Frontiers in psychology.* 2012;3:234.
11. O'Haire ME, McKenzie SJ, Beck AM, Slaughter V. Animals may act as social buffers: Skin conductance arousal in children with autism spectrum disorder in a social context. *Dev Psychobiol.* 2015;57(5):584-595.
12. Christon LM, Mackintosh VH, Myers BJ. Use of complementary and alternative medicine (CAM) treatments by parents of children with autism spectrum disorders. *Research in Autism Spectrum Disorders.* 2010;4(2):249-259.
13. Gabriels RL, Pan Z, Dechant B, Agnew JA, Brim N, Mesibov G. Randomized Controlled Trial of Therapeutic Horseback Riding in Children and Adolescents With Autism Spectrum Disorder. *J Am Acad Child Adolesc Psychiatry.* 2015;54(7):541-549.
14. Gabriels RL, Pan Z, Guerin NA, Dechant B, Mesibov G. Six- month outcomes of therapeutic horseback riding in children and adolescents with autism spectrum disorder. In: University of Colorado Anschutz Medical Campus.
15. Gabriels RL, Agnew JA, Holt KD, et al. Pilot study measuring the effects of therapeutic horseback riding on school-age children and adolescents with autism spectum disorders. *Research in Autism Spectrum Disorders.* 2012;6:578 - 588.
16. Shoffner A, Gabriels R. Therapeutic Horseback Riding Intervention Manual. In. Denver, CO: University of Colorrado and Children's Hosptial Colorado; 2008.
17. PATH International. PATH International. http://www.pathintl.org/. Published 2011. Accessed 02/21/2018.
18. Pan Z, Granger DA, Guerin NA, Shoffner A, Gabriels RL. Replication Pilot Trial of Therapeutic Horseback Riding and Cortisol Collection With Children on the Autism Spectrum. *Frontiers in veterinary science.* 2018;5:312.
19. Leyfer OT, Folstein SE, Bacalman S, et al. Comorbid psychiatric disorders in children with autism: Interview development and rates of disorders. *Journal of autism and developmental disorders.* 2006.
20. Joshi G, Petty C, Wozniak J, et al. The heavy burden of psychiatric comorbidity in youth with autism spectrum disorders: a large comparative study of a psychiatrically referred population. *Journal of autism and developmental disorders.* 2010;40(11):1361-1370.
21. Mayes SD, Calhoun SL, Murray MJ, Ahuja M, Smith LA. Anxiety, depression, and irritability in children with autism relative to other neuropsychiatric disorders and typical development. *Research in Autism Spectrum Disorders.* 2011;5(1):474-485.
22. Szatmari P, McConnell B. Anxiety and mood disorders in individuals with autism spectrum disorder. *Autism spectrum disorders.* 2011:330-338.

23. van Steensel FJ, Bögels SM, Perrin S. Anxiety disorders in children and adolescents with autistic spectrum disorders: a meta-analysis. *Clinical child and family psychology review.* 2011;14(3):302.
24. Doehring P, Reichow B, Palka T, Phillips C, Hagopian L. Behavioral approaches to managing severe problem behaviors in children with autism spectrum and related developmental disorders: a descriptive analysis. *Child and Adolescent Psychiatric Clinics.* 2014;23(1):25-40.
25. Siegel M, Doyle K, Chemelski B, et al. Specialized inpatient psychiatry units for children with autism and developmental disorders: a United States survey. *Journal of autism and developmental disorders.* 2012;42(9):1863-1869.
26. Argumedes M, Lanovaz MJ, Larivée S. Brief Report: Impact of Challenging Behavior on Parenting Stress in Mothers and Fathers of Children with Autism Spectrum Disorders. *Journal of autism and developmental disorders.* 2018:1-5.
27. American Psychiatric A, American Psychiatric Association DSMTF, Joint Information Service of the American Psychiatric A, et al. *Diagnostic and statistical manual of mental disorders DSM-5.* 5th ed.. ed. Arlington, VA: Arlington, VA : American Psychiatric Association; 2013.
28. Weitlauf AS, McPheeters ML, Peters B, et al. Therapies for children with autism spectrum disorder. 2014.
29. Croen LA, Najjar DV, Ray GT, Lotspeich L, Bernal P. A comparison of health care utilization and costs of children with and without autism spectrum disorders in a large group-model health plan. *Pediatrics.* 2006;118(4):e1203-1211.
30. Pedersen KA, Santangelo SL, Gabriels RL, Righi G, Erard M, Siegel M. Behavioral Outcomes of Specialized Psychiatric Hospitalization in the Autism Inpatient Collection (AIC): A Multisite Comparison. *Journal of autism and developmental disorders.* 2017:1-10.
31. Aman MG, Burrow WH, Wolford PL. The Aberrant Behavior Checklist-Community: Factor validity and effect of subject variables for adults in group homes. *Am J Ment Retard.* 1995;100(3):283-292.
32. McCracken JT, McGough J, Shah B, et al. Risperidone in children with autism and serious behavioral problems. *N Engl J Med.* 2002;347(5):314-321.
33. Owen R, Sikich L, Marcus RN, et al. Aripiprazole in the treatment of irritability in children and adolescents with autistic disorder. *Pediatrics.* 2009;124(6):1533-1540.
34. Mazefsky C, Kao J, Oswald D. Preliminary evidence suggesting caution in the use of psychiatric self-report measures with adolescents with high-functioning autism spectrum disorders. *Research in Autism Spectrum Disorders.* 2011;5(1):164-174.
35. Nayfack AM, Huffman LC, Feldman HM, Chan J, Saynina O, Wise PH. Hospitalizations of children with autism increased from 1999 to 2009. *Journal of autism and developmental disorders.* 2014;44(5):1087-1094.
36. Thompson RA. Emotion regulation: A theme in search of definition. *Monographs of the society for research in child development.* 1994;59(2-3):25-52.
37. Aldao A, Nolen-Hoeksema S, Schweizer S. Emotion-regulation strategies across psychopathology: A meta-analytic review. *Clin Psychol Rev.* 2010;30(2):217-237.
38. Mazefsky CA. Emotion regulation and emotional distress in autism spectrum disorder: Foundations and considerations for future research. In: Springer; 2015.
39. Mazefsky CA, Herrington J, Siegel M, et al. The role of emotion regulation in autism spectrum disorder. *Journal of the American Academy of Child & Adolescent Psychiatry.* 2013;52(7):679-688.
40. Bouscein W. Electrodermal Activity. 2012.
41. Cook JE, Leventhal B, Heller W, Metz J, Wainwright M, Freedman D. Autistic children and their first-degree relatives: relationships between serotonin and norepinephrine levels and intelligence. *The Journal of neuropsychiatry and clinical neurosciences.* 1990;2(3):268-274.
42. Gillberg C, Svennerholm L. CSF monoamines in autistic syndromes and other pervasive developmental disorders of early childhood. *The British Journal of Psychiatry.* 1987;151(1):89-94.
43. Lake CR, Ziegler MG, Murphy DL. Increased norepinephrine levels and decreased dopamine-β-hydroxylase activity in primary autism. *Arch Gen Psychiatry.* 1977;34(5):553-556.
44. Launay J-M, Bursztejn C, Ferrari P, et al. Catecholamines metabolism in infantile autism: a controlled study of 22 autistic children. *Journal of autism and developmental disorders.* 1987;17(3):333-347.
45. Leboyer M, Bouvard MP, Launay J-M, et al. Brief report: A double-blind study of naltrexone in infantile autism. *Journal of autism and developmental disorders.* 1992;22(2):309-319.
46. Leventhal BL, Cook EH, Morford M, Ravitz A, Freedman DX. Relationships of whole blood serotonin and plasma norepinephrine within families. *Journal of autism and developmental disorders.* 1990;20(4):499-511.
47. Schoen SA, Miller LJ, Brett-Green BA, Nielsen DM. Physiological and Behavioral Differences in Sensory Processing: A Comparison of Children with Autism Spectrum Disorder and Sensory Modulation Disorder. *Frontiers in Integrative Neuroscience.* 2009;3:29.
48. Schaaf RC, Benevides TW, Leiby BE, Sendecki JA. Autonomic dysregulation during sensory stimulation in children with autism spectrum disorder. *Journal of autism and developmental disorders.* 2015;45(2):461-472.
49. Woodard CR, Goodwin MS, Zelazo PR, et al. A comparison of autonomic, behavioral, and parent-report measures of sensory sensitivity in young children with autism. *Research in Autism Spectrum Disorders.* 2012;6(3):1234-1246.
50. Lydon S, Healy O, Reed P, Mulhern T, Hughes BM, Goodwin MS. A systematic review of physiological reactivity to stimuli in autism. *Dev Neurorehabil.* 2016;19(6):335-355.
51. Levine TP, Sheinkopf SJ, Pescosolido M, Rodino A, Elia G, Lester B. Physiologic arousal to social stress in children with autism spectrum disorders: a pilot study. *Research in autism spectrum disorders.* 2012;6(1):177-183.

52. Ming X, Julu PO, Brimacombe M, Connor S, Daniels ML. Reduced cardiac parasympathetic activity in children with autism. *Brain Dev.* 2005;27(7):509-516.
53. Porges SW, Macellaio M, Stanfill SD, et al. Respiratory sinus arrhythmia and auditory processing in autism: Modifiable deficits of an integrated social engagement system? *International Journal of Psychophysiology.* 2013;88(3):261-270.
54. Toichi M, Kamio Y. Paradoxical autonomic response to mental tasks in autism. *Journal of autism and developmental disorders.* 2003;33(4):417-426.
55. Zamzow RM, Ferguson BJ, Stichter JP, et al. Effects of propranolol on conversational reciprocity in autism spectrum disorder: a pilot, double-blind, single-dose psychopharmacological challenge study. *Psychopharmacology.* 2016;233(7):1171-1178.
56. Anderson CJ, Colombo J, Unruh KE. Pupil and salivary indicators of autonomic dysfunction in autism spectrum disorder. *Dev Psychobiol.* 2013;55(5):465-482.
57. Corbett BA, Mendoza S, Abdullah M, Wegelin JA, Levine S. Cortisol circadian rhythms and response to stress in children with autism. *Psychoneuroendocrinology.* 2006;31(1):59-68.
58. Spratt EG, Nicholas JS, Brady KT, et al. Enhanced cortisol response to stress in children in autism. *Journal of autism and developmental disorders.* 2012;42(1):75-81.
59. Taylor JL, Corbett BA. A review of rhythm and responsiveness of cortisol in individuals with autism spectrum disorders. *Psychoneuroendocrinology.* 2014;49:207-228.
60. Mikita N, Hollocks MJ, Papadopoulos AS, et al. Irritability in boys with autism spectrum disorders: an investigation of physiological reactivity. *J Child Psychol Psychiatry.* 2015;56(10):1118-1126.
61. Levine TP, Conradt E, Goodwin MS, Sheinkopf SJ, Lester B. Psychophysiological arousal to social stress in autism spectrum disorders. In: *Comprehensive guide to autism.* Springer; 2014:1177-1193.
62. Coury DL, Anagnostou E, Manning-Courtney P, et al. Use of psychotropic medication in children and adolescents with autism spectrum disorders. *Pediatrics.* 2012;130(Supplement 2):S69-S76.
63. Marcus RN, Owen R, Manos G, et al. Safety and tolerability of aripiprazole for irritability in pediatric patients with autistic disorder: a 52-week, open-label, multicenter study. *The Journal of clinical psychiatry.* 2011;72(9):1270-1276.
64. Arnold LE. The alone rangers and silver. *Journal of the American Academy of Child & Adolescent Psychiatry.* 2015;54(7):535-536.
65. Risperidone Statistical PREA - FDA. U.S. Department of Health and Human Services, Food and Dtug Administration. https://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/DevelopmentResources/UCM322383.pdf. Published 2011. Accessed 03/25/2019, 2019.
66. Ung D, Selles R, Small BJ, Storch EA. A systematic review and meta-analysis of cognitive-behavioral therapy for anxiety in youth with high-functioning autism spectrum disorders. *Child Psychiatry & Human Development.* 2015;46(4):533-547.
67. de Bruin EI, Blom R, Smit FM, van Steensel FJ, Bögels SM. MYmind: Mindfulness training for youngsters with autism spectrum disorders and their parents. *Autism.* 2015;19(8):906-914.
68. Virués-Ortega J. Applied behavior analytic intervention for autism in early childhood: Meta-analysis, meta-regression and dose–response meta-analysis of multiple outcomes. *Clin Psychol Rev.* 2010;30(4):387-399.
69. Reichow B. Overview of meta-analyses on early intensive behavioral intervention for young children with autism spectrum disorders. *Journal of autism and developmental disorders.* 2012;42(4):512-520.
70. Lerner MD, White SW, McPartland JC. Mechanisms of change in psychosocial interventions for autism spectrum disorders. *Dialogues Clin Neurosci.* 2012;14(3):307-318.
71. O'Haire ME. Research on animal-assisted intervention and autism spectrum disorder, 2012–2015. *Applied Developmental Science.* 2017;21(3):200-216.
72. Peters BCM, Wood W. Autism and Equine-Assisted Interventions: A Systematic Mapping Review. *Journal of autism and developmental disorders.* 2017;47(10):3220-3242.
73. Srinivasan SM, Cavagnino DT, Bhat AN. Effects of Equine Therapy on Individuals with Autism Spectrum Disorder: A Systematic Review. *Rev J Autism Dev Disord.* 2018;5(2):156-175.
74. Gabriels RL, Pan Z, Guerin NA, Dechant B, Mesibov G. Long-Term Effect of Therapeutic Horseback Riding in Youth With Autism Spectrum Disorder: A Randomized Trial. *Frontiers in veterinary science.* 2018;5:156.
75. Anestis MD, Anestis JC, Zawilinski LL, Hopkins TA, Lilienfeld SO. Equine-related treatments for mental disorders lack empirical support: A systematic review of empirical investigations. *Journal of clinical psychology.* 2014;70(12):1115-1132.
76. Selby A, Smith-Osborne A. A systematic review of effectiveness of complementary and adjunct therapies and interventions involving equines. *Health Psychol.* 2013;32(4):418-432.
77. Schretzmayer L, Kotrschal K, Beetz A. Minor Immediate Effects of a Dog on Children's Reading Performance and Physiology. *Frontiers in veterinary science.* 2017;4:90.
78. Viau R, Arsenault-Lapierre G, Fecteau S, Champagne N, Walker C-D, Lupien S. Effect of service dogs on salivary cortisol secretion in autistic children. *Psychoneuroendocrinology.* 2010;35:1187 - 1193.
79. Pendry P, Smith AN, Roeter SM. Randomized trial examines effects of equine facilitated learning on adolescents' basal cortisol levels. *Human-Animal Interaction Bulletin.* 2014;2(1):80-95.
80. Tabares C, Vicente F, Sánchez S, Aparicio A, Alejo S, Cubero J. Quantification of hormonal changes by effects of hippotherapy in the autistic population. *Neurochemical Journal.* 2012;6(4):311-316.

81. Thayer JF, Hahn AW, Pearson MA, Sollers JJ, 3rd, Johnson PJ, Loch WE. Heart rate variability during exercise in the horse. *Biomed Sci Instrum.* 1997;34:246-251.
82. Carter JB, Banister EW, Blaber AP. The effect of age and gender on heart rate variability after endurance training. *Med Sci Sports Exerc.* 2003;35(8):1333-1340.
83. Levy WC, Cerqueira MD, Harp GD, et al. Effect of endurance exercise training on heart rate variability at rest in healthy young and older men. *Am J Cardiol.* 1998;82(10):1236-1241.
84. Tonello L, Rodrigues FB, Souza JW, Campbell CS, Leicht AS, Boullosa DA. The role of physical activity and heart rate variability for the control of work related stress. *Front Physiol.* 2014;5:67.
85. Bishop SR, Lau M, Shapiro S, et al. Mindfulness: A proposed operational definition. *Clinical psychology: Science and practice.* 2004;11(3):230-241.
86. Yerkes RM, Dodson JD. The relation of strength of stimulus to rapidity of habit-formation. *Journal of comparative neurology.* 1908;18(5):459-482.
87. Crossman MK. Effects of Interactions With Animals On Human Psychological Distress. *Journal of clinical psychology.* 2017;73(7):761-784.
88. Kazdin AE. Mediators and mechanisms of change in psychotherapy research. *Annu Rev Clin Psychol.* 2007;3:1-27.
89. Schreibman L, Dufek S, Cunningham AB. Identifying moderators of treatment outcome for children with autism. In: *International handbook of autism and pervasive developmental disorders.* Springer; 2011:295-305.
90. MacKinnon DP, Lockhart G, Baraldi AN, Gelfand LA. 15 Evaluating Treatment Mediators and Moderators. *The Oxford handbook of research strategies for clinical psychology.* 2013:262.
91. Arnold LE, Aman MG, Martin A, et al. Assessment in multisite randomized clinical trials of patients with autistic disorder: The Autism RUPP Network. Research Units on Pediatric Psychopharmacology. *Journal of autism and developmental disorders.* 2000;30(2):99-111.
92. MacKinnon DP. Integrating mediators and moderators in research design. *Research on social work practice.* 2011;21(6):675-681.
93. Health NIoM. Arousal and Regulatory Systems: Workshop Proceedings. National Institute of Mental Health. https://www.nimh.nih.gov/research-priorities/rdoc/arousal-and-regulatory-systems-workshop-proceedings.shtml. Published 2012. Accessed 03/12/2018, 2018.
94. Dardas LA, Ahmad MM. Validation of the World Health Organization's quality of life questionnaire with parents of children with autistic disorder. *Journal of autism and developmental disorders.* 2014;44(9):2257-2263.
95. Group W. Development of the World Health Organization WHOQOL-BREF quality of life assessment. *Psychol Med.* 1998;28(3):551-558.
96. Ward SC, Whalon K, Rusnak K, Wendell K, Paschall N. The association between therapeutic horseback riding and the social communication and sensory reactions of children with autism. *Journal of autism and developmental disorders.* 2013;43(9):2190-2198.
97. Smith T, Scahill L, Dawson G, et al. Designing research studies on psychosocial interventions in autism. *Journal of autism and developmental disorders.* 2007;37(2):354-366.
98. Aman MG, Singh NN, Stewart AW, Field CJ. The aberrant behavior checklist: A behavior rating scale for the assessment of treatment effects. *Am J Ment Defic.* 1985;89(5):485-491.
99. Zhaoxing P, Shoffner A, Guerin NA, Gabriels RL. Replication pilot trial of therapeutic horseback riding and cortisol collection with children on the autism spectrum. In: University of Colorado Anschutz Medical Campus; 2017.
100. Inc. E. Empatica E4 Wristband. https://www.empatica.com/en-eu/research/e4/. Published 2018. Accessed 03/12/2018, 2018.
101. Germone MM, Gabriels RL, Guerin NA, Pan Z, Banks T, O'Haire ME. Animal-assisted activity improves social behaviors in psychiatrically hospitalized youth with autism. *Autism.* 2019:1362361319827411.
102. Heathers JA, Fink E, Kuhnert RL, de Rosnay M. Blood volume pulse (BVP) derived vagal tone (VT) between 5 and 7 years of age: a methodological investigation of measurement and longitudinal stability. *Dev Psychobiol.* 2014;56(1):23-35.
103. Liss M, Mailloux J, Erchull MJ. The relationships between sensory processing sensitivity, alexithymia, autism, depression, and anxiety. *Personality and Individual Differences.* 2008;45(3):255-259.
104. Gaigg SB, Cornell AS, Bird G. The psychophysiological mechanisms of alexithymia in autism spectrum disorder. *Autism.* 2016.
105. Bird G, Cook R. Mixed emotions: the contribution of alexithymia to the emotional symptoms of autism. *Translational Psychiatry.* 2013;3:e285.
106. Roid GH, Miller LJ. *Leiter International Performance Scale-Revised (Leiter-R).* Wood Dale, IL: Stoelting; 1997.
107. Gadow KD, Sprafkin J. *Child and Adolescent Symptom Inventory.* Stony Brook, NY: Checkmate Plus; 2005.
108. Rutter M, Bailey A, Lord C. *Social Communication Questionnaire.* Los Angeles, CA: Western Psychological Services; 2003.
109. Norris M, Lecavalier L. Screening accuracy of level 2 autism spectrum disorder rating scales: A review of selected instruments. *Autism.* 2010;14(4):263-284.
110. Lord C, Rutter M, diLavore PC, Risi S, Gotham K, Bishop S. *Autism Diagnostic Observation Schedule 2nd Edition Manual.* Los Angeles: Western Psychological Services; 2012.
111. Granger DA, Johnson SB, Szanton SL, Out D, Schumann LL. Incorporating salivary biomarkers into nursing research: an overview and review of best practices. *Biological research for nursing.* 2012;14(4):347-356.
112. Guerin NA, Gabriels RL, Germone MM, et al. Reliability and Validity Assessment of the Observation of Human-Animal Interaction for Research (OHAIRE) Behavior Coding Tool. *Frontiers in veterinary science.* 2018;5:268.
113. Gabriels RL, Ivers BJ, Hill DE, Agnew JA, McNeill J. Stability of adaptive behaviors in middle-school children with autism spectrum disorders. *Research in Autism Spectrum Disorders.* 2007;1(4):291-303.

114. Nam CB, Boyd M. Occupational status in 2000; over a century of census-based measurement. *Population Research and Policy Review.* 2004;23(4):327-358.
115. Lavigne JV, Cromley T, Sprafkin J, Gadow KD. The child and adolescent symptom inventory-progress monitor: a brief diagnostic and statistical manual of mental disorders, -referenced parent-report scale for children and adolescents. *J Child Adolesc Psychopharmacol.* 2009;19(3):241-252.
116. Gadow KD, Sprafkin J, Carlson GA, et al. A DSM-IV–referenced, adolescent self-report rating scale. *Journal of the American Academy of Child & Adolescent Psychiatry.* 2002;41(6):671-679.
117. Lecavalier L, Gadow KD, DeVincent CJ, Edwards MC. Validation of DSM-IV model of psychiatric syndromes in children with autism spectrum disorders. *Journal of autism and developmental disorders.* 2009;39(2):278-289.
118. Gadow KD, Sprafkin J. The symptom inventories: An annotated bibliography. *Stony Brook, NY: Checkmate Plus.* 2009.
119. Clow A, Hucklebridge F, Stalder T, Evans P, Thorn L. The cortisol awakening response: more than a measure of HPA axis function. *Neurosci Biobehav Rev.* 2010;35(1):97-103.
120. Corbett BA, Mendoza S, Wegelin JA, Carmean V, Levine S. Variable cortisol circadian rhythms in children with autism and anticipatory stress. *J Psychiatry Neurosci.* 2008;33(3):227-234.
121. Sharpley CF, Bitsika V, Andronicos NM, Agnew LL. Is afternoon cortisol more reliable than waking cortisol in association studies of children with an ASD? *Physiology & behavior.* 2016;155:218-223.
122. van Dooren M, de Vries JJG, Janssen JH. Emotional sweating across the body: Comparing 16 different skin conductance measurement locations. *Physiology & Behavior.* 2012;106(2):298-304.
123. Pieper JR, Laugero KD. Preschool children with lower executive function may be more vulnerable to emotional-based eating in the absence of hunger. *Appetite.* 2013;62:103-109.
124. O'Haire ME, McKenzie SJ, Beck AM, Slaughter V. Animals may act as social buffers: Skin conductance arousal in children with autism spectrum disorder in a social context. *Dev Psychobiol.* 2015;57(5):584-595.
125. Ellis RJ, Zhu B, Koenig J, Thayer J, Wang Y. A careful look at ECG sampling frequency and R-peak interpolation on short-term measures of heart rate variability. *Physiological Measurement.* 2015;36(9):1827.
126. Constantino JN. *The Social Responsiveness Scale.* Los Angeles: Western Psychological Services; 2002.
127. *SALT: A computer program for the Systematic Analysis of Language Transcripts* [computer program]. Madison, WI: University of Wisconsin; 2000.
128. Mazefsky CA, Day TN, Siegel M, White SW, Yu L, Pilkonis PA. Development of the Emotion Dysregulation Inventory: A PROMIS(R)ing Method for Creating Sensitive and Unbiased Questionnaires for Autism Spectrum Disorder. *Journal of autism and developmental disorders.* 2016.
129. Mazefsky CA, Yu L, White SW, Siegel M, Pilkonis PA. The Emotion Dysregulation Inventory: Psychometric properties and item response theory calibration in an autism spectrum disorder sample. In. *Autism Research*2018.
130. Hambleton RK, Swaminathan H, Rogers HJ. *Fundamentals of item response theory.* Vol 2: Sage; 1991.
131. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap) - A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics.* 2009;42(2):377 - 381.
132. Kleckner IR, Jones RM, Wilder-Smith O, et al. Simple, Transparent, and Flexible Automated Quality Assessment Procedures for Ambulatory Electrodermal Activity Data. *IEEE Trans Biomed Eng.* 2017.
133. Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *Eur Heart J.* 1996;17(3):354-381.
134. Scholkmann F, Boss J, Wolf M. An efficient algorithm for automatic peak detection in noisy periodic and quasi-periodic signals. *Algorithms.* 2012;5(4):588-603.
135. Quintana DS, Alvares GA, Heathers JA. Guidelines for Reporting Articles on Psychiatry and Heart rate variability (GRAPH): recommendations to advance research communication. *Transl Psychiatry.* 2016;6:e803.
136. Quintana DS, Heathers JA. Considerations in the assessment of heart rate variability in biobehavioral research. *Frontiers in psychology.* 2014;5:805.
137. Kelsey M, Akcakaya M, Kleckner IR, et al. Applications of sparse recovery and dictionary learning to enhance analysis of ambulatory electrodermal activity data. *Biomedical Signal Processing and Control.* 2018;40:58-70.
138. MacKinnon DP, Lockwood CM, Hoffman JM, West SG, Sheets V. A comparison of methods to test mediation and other intervening variable effects. *Psychological methods.* 2002;7(1):83.
139. Fritz MS, MacKinnon DP. Required sample size to detect the mediated effect. *Psychol Sci.* 2007;18(3):233-239.
140. Bryson SE. Brief report: Epidemiology of autism. *Journal of autism and developmental disorders.* 1996;26(2):165-167.
141. Gabriels RL, Agnew JA, Holt KD, et al. Pilot study measuring the effects of therapeutic horseback riding on school-age children and adolescents with autism spectrum disorders. *Research in Autism Spectrum Disorders.* 2012;6(2):578-588.
142. Brain CfAatD. General Questions: Training, Reliability, Becoming a Trainer, Translation. Weill Cornell Medicine Psychiatry. https://psychiatry.weill.cornell.edu/sites/default/files/faqs.pdf. Published 2013. Accessed 03/12/2018, 2018.