THE ROLE OF FRUSTRATION IN INTENSIVE TREATMENT OF CHILDHOOD APRAXIA OF SPEECH

A Thesis
Submitted to
the Temple University Graduate Board

In Partial Fulfillment
Of the Requirements for the Degree
MASTER OF ARTS
OF SPEECH, LANGUAGE, AND HEARING SCIENCE

---

by
Shina J. Cook
Diploma Date May 2022

Examining Committee Members:

Edwin Maas, Thesis Advisor, Communication Sciences and Disorders
Susan Caspari, Communication Sciences and Disorders
Rebecca Alper, Communication Sciences and Disorders

ABSTRACT

Purpose: This study primarily investigates the effects and influence of frustration in children with childhood apraxia of speech (CAS) in the setting of intensive treatment. Additionally, the study examines the interrater reliability of the frustration rating scale used in an intensive CAS treatment study.

Methods: Frustration and treatment data obtained from 17 participants (between 4;0-9;11 years) with CAS in an intensive treatment research study were retrospectively used to determine potential relationships related to frustration in treatment (target complexity, temporal conditions, session number, CAS severity). Interrater reliability of the frustration rating scale was assessed with 34 randomly selected treatment session videos scored by a blinded second rater and compared to original scores.

Results: Interrater reliability of the scale was poor to fair but had relatively close agreement within one scale point. Frustration levels were observed to decrease over the course of the treatment period but were typically greater in the afternoon sessions compared to morning. Participants in the complex target treatment condition with lower frustration also exhibited better outcomes than those with greater frustration. No other relationships were observed.

Conclusions: Due to relatively poor interrater reliability of frustration scoring system used to obtain data used in the current study, results of the study should be interpreted cautiously. There may be a relationship between frustration levels in children with CAS and treatment conditions and outcomes, but other factors may influence both variables and further investigation into frustration is necessary to draw stronger conclusions.

ACKNOWLEDGEMENTS

I would like to extend my thanks to the following important presences throughout the course of this project:

First and foremost, I would like to thank Edwin Maas for being a highly dedicated and inspiring thesis advisor. This project would not have been possible without his words of encouragement, continued support, and willingness to explain basic statistics to me.

Next, Sue Caspari and Rebecca Alper, the two remaining members of my thesis committee, contributed their time, support, and wealth of knowledge to my project, and I owe them a tremendous amount of gratitude. Additionally, I would like to thank Sue Caspari for introducing me to the world of childhood apraxia of speech early in my graduate studies and fostering my continued interest throughout my education.

Further, I thank the participants, clinicians, and all those behind the scenes involved in the ASSIST treatment study – without them, this project would not exist.

I also wish to thank Lauren Colletti for the long nights, the charcuterie boards, and for sharing the journey of writing a Master's thesis with me. Throughout this project, she was a rock when I need to stay grounded, ears when I needed to complain, and a thesaurus when I couldn't think of the right adjective.

To my clinical supervisors, Allison Snyder and Tatiana Bereznikova, thank you for teaching me what it means to be a Speech-Language Pathologist in the real world.

Finally, to my parents, Tori, Rachel, Kelsey, Dana, Vincenzo, Gigi, Tonton, the green Halal cart on Temple's campus, thank you for always being there and supporting me throughout this process.

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER 1

INTRODUCTON

Background

Childhood apraxia of speech (CAS) is a pediatric speech sound disorder characterized by an impairment of the planning and programming of movement sequences required for speech, resulting in errored speech sounds and prosody, not explainable by neuromuscular deficits (ASHA, 2007). CAS has been reported to negatively impact academic performance, social-emotional development, and reading outcomes (Miller et al., 2019).

Due to limited research, the exact prevalence of CAS is unknown. It is estimated that one to ten children out of 1000 (between 40,000-400,000 total) are affected in the United States (Shriberg et al., 1997). According to the ASHA 2020 School Survey (ASHA, 2020), 57.6% of school-based speech-language pathologists (SLP) have at least one child with CAS on their caseload, 2.7 per SLP on average. Although over half of the school SLPs work with at least one child with CAS, the disorder is not adequately addressed in most graduate programs, leaving many clinicians underprepared to treat it (Apraxia Kids, 2021). Similar to other neurobehavioral disorders, such as autism spectrum disorder (ASD) and attention deficit hyperactivity disorder (ADHD), CAS diagnoses have rapidly increased in the last couple of decades due to expanded research and awareness (ASHA, 2007). With the increased number of diagnosed children, there is also an influx of children with CAS who need specialized treatment for their motor

speech disorder; however, the current need for treatment exceeds the readiness and availability of the treatment system.

CAS treatments are traditionally recommended to be intensive (ASHA, 2007), with frequent sessions and rigorous drill-based therapy as dictated by the principles of motor learning (Davis & Velleman, 2000; Namasivayam et al., 2015), which may ultimately take a mental and emotional toll on the child. Several intensive treatments have been developed for CAS, such as Dynamic Temporal and Tactile Cueing (DTTC; Strand, 2020; Strand et al., 2006), Motor Speech Treatment Protocol (Namasivayam et al., 2015), Rapid Syllable Transition Training (ReST; Ballard et al., 2010; Murray et al., 2015), and Apraxia of Speech Systematic Integral Stimulation Treatment (ASSIST; Maas et al., 2019). Each of these treatment protocols have variable degrees of evidence in support of their efficacy, yet there has been no effort to examine the emotional ramifications of these intensive treatment strategies upon the receiving child.

Heightened frustration is an example of a negative emotional impact of intensive treatment. There has been to no research regarding the role of frustration in achieving the desired treatment intensity or gains. Clinicians need information about treatment benefits and risks (e.g., frustration) to make evidence-based decisions for their clients. Thus, systematic investigation of frustration associated with treatment parameters, such as treatment target choices or time of day, and its relation to treatment gains is warranted.

Frustration

Frustration, as defined in the dictionary, refers to "a deep chronic sense or state of insecurity and dissatisfaction arising from unresolved problems or unfulfilled needs" (Merriam Webster, 2021). The definition of "frustrating", according to the same dictionary, is "causing feelings of anger and annoyance" (Merriam Webster, 2021). Frustration is a difficult feeling to quantify, as it may be felt and expressed differently from person to person. Some people may experience a physiological reaction, such as an increased heart rate or sweaty hands, while for others, it may be a purely psychological response. In the research literature, frustration (or rather, frustration tolerance) is often operationalized by measuring the time taken before a participant abandons an impossible task (e.g., Rardin & Moan, 1970), percent of time spent on a difficult (near-impossible) task (Meindl et al., 2019), or via self-reported questionnaires (e.g., Longo et al., 2016).

Children with speech sound disorders (SSD), such as CAS, demonstrate difficulties communicating wants and needs due to reduced intelligibility. These challenges are examples of unresolved problems and unfulfilled goals and can therefore be expected to result in frustration. Indeed, frustration has been reported in children with SSD (e.g., McCormack et al., 2010; Rusiewicz, Maize, & Ptakowski, 2018). Around 22.2% of parents of a child with CAS reported that their child felt frustrated due to their speech (Rusiewicz et al., 2018). There are also reports of children with SSD experiencing frustration related to their inability to "speak properly" and their communication partner's inability to "listen properly" (McCormack et al., 2010). Boys are more likely to experience negative effects of frustration on their performance (Solkoff et al., 1964), and

to have SSD (Adani & Cepanec, 2019; Wren, Roulstone, & Miller, 2012), including CAS (Hall, Jordan, & Robin, 1993; Lewis et al., 2004; Shriberg, Kwiatkowski, & Mabie, 2019). This suggests that children with CAS, since they are typically boys, are at increased risk of negative effects of frustration.

It is reasonable to assume that participating in tasks that are intense, difficult, strenuous, and/or repetitive increases frustration (e.g., Meindl et al., 2019). Speech therapy for CAS often meets these criteria (intense, difficult, repetitive), and therefore may further increase frustration in children with CAS, who are already at risk of increased frustration. Consequently, factors related to treatment (e.g., complexity of targets; number of sessions) may affect frustration as a dependent variable.

Frustration can also be considered an independent variable, in that it can impact a person's performance on a task (Solfkoff et al., 1964). Although there are no studies that focus directly on frustration in children during speech therapy, there is evidence to suggest that college aged adults with higher levels of anxiety performed at a detriment on complex intellectual tasks following a task designed to increase their frustration compared to participants with lower levels of anxiety (Hokanson & Burgess, 1964). Similarly, greater frustration tolerance predicts higher grade point average and success toward a college degree (Meindl et al., 2019; Rardin & Moan, 1970). Since parents of children with CAS have also reported high rates of frustration in their children (Rusiewicz et al., 2018), it could be argued that frustration levels may also impact a child's performance on difficult tasks, such as treatment. This impact could be positive or negative, and both should be considered. This sets up the question of differences in

outcomes among children with CAS following intensive treatment tasks with higher versus lower rates of frustration.

Internal factors of the child including temperament, disposition, and self-awareness may also influence frustration in treatment. The child's baseline for frustration tolerance and overall personality may be a factor that relates to their potential to experience frustration. Grit, defined as "the tenacious pursuit of a dominant superordinate goal despite setbacks" (Duckworth & Gross, 2014), is an internal characteristic to consider in the discussion of frustration in the face of difficult tasks. As mentioned above, children with CAS face several obstacles related to their communication, including intensive treatment. Frustration exhibited in treatment may be influenced by internal factors such as grit, as "gritty individuals" may be more capable of suppressing behaviors that would be perceived as frustration (e.g., crying, noncompliance) in pursuit of the higher goal (e.g., successful outcomes in treatment) (Duckworth & Gross, 2014). The child's awareness of their communicative difficulties and the purpose of speech therapy would be a key factor in their ability to conceive such "higher goals"; children less aware of their speech inaccuracies or the objective of treatment may be less motivated to suppress behaviors in pursuit of improvement.

In sum, since CAS is a potential risk factor for increased frustration, it is important to identify treatment methods and conditions that are related to fewer negative impacts of frustration in children to promote better performance and treatment progress.

Present Study

The present study is the first to systematically investigate frustration in relation to speech therapy in children with CAS. Specifically, this study will examine frustration in the context of a Phase 1 randomized controlled trial study of Apraxia of Speech Systematic Integral Stimulation Treatment (ASSIST), an intensive drill-based treatment for CAS based on integral stimulation and principles of motor learning. Initial evidence from a small-scale single-case experimental design study (n=6) suggests that a less intensive version of ASSIST resulted in improved speech production for children with CAS (Maas et al., 2019). However, the present study of ASSIST involves a more intensive implementation, as it was run as a summer camp during which children received 16 hours of ASSIST in a 2-week period (2 hours per day, divided over 4 sessions of 30 minutes). As such, this implementation runs the risk of increased frustration in children. As a Phase 1 study, the goals are to determine initial efficacy (to be reported elsewhere) and to evaluate safety by considering negative side effects (in this case, frustration). The focus of the present study is on the latter.

Frustration levels of each child were taken after each treatment session, to monitor for the wellbeing of the child and take appropriate action as needed (e.g., change reinforcers; discontinue treatment). Frustration was rated by the clinician performing the treatment and defined by the judgement of the clinician. Several considerations led to this approach to measurement of frustration. First, a measure was needed that could provide information about frustration quickly, on a session-by-session basis, to inform immediate decisions about changes or discontinuation of treatment.

Second, the measure would need to account for possible differences in how frustration is expressed in different children. Although children's self-report of frustration may be considered to provide the most direct insight, such self-report measures are unlikely to be valid or reliable considering the complex concept of frustration and the age range of the children (4 to 9 years old).

Third, frustration may be expressed differently by different children (e.g., some children may cry, others may simply withdraw or refuse to cooperate), necessitating some measure that could take such differences into account. Thus, counting instances of crying (for example), would likely fail to capture frustration in all children.

Fourth, the measurement of frustration should be feasible in a clinical setting. While physiological measurements (e.g. heart rate, galvanic skin response) may be more objective than clinician judgement, such measures require instrumentation and more extensive data processing. Therefore, judgement by the treating clinician was considered a suitable measure for this initial investigation.

However, since frustration levels are being determined based on subjective judgement, it raises the question of reliability. The present study will first address this question by randomly selecting a subset of sessions to be evaluated for frustration by an independent clinician utilizing the same rating scale. Scores will be compared to determine the inter-rater reliability of the rating scale system.

Positive effects of frustration in the setting of intensive treatment may present as increased grit or perseverance to complete tasks, which could result in improved performance on aforementioned tasks. Negative impacts of frustration may include avoidance of task by focusing attention elsewhere or through behaviors that present as

barriers to treatment, such as tearfulness or physical abandonment of task (e.g., walking away from table). The present study utilizes a behavioral measure of frustration; scores are determined based on subjective interpretations of behaviors that are typically deemed as negative effects. While it is possible that positive effects of frustration exist, they may not be addressed in the setting of the present study due to the framework in which the frustration data is collected as negative effects are the sole determinants for frustration scores. Negative effects are both simpler to observe (i.e., outward behaviors that delay treatment vs. internal feelings influencing motivation) and more related to participant safety (i.e., determining emotional wellbeing of children during treatment), which was the primary purpose of collecting frustration data in the setting of ASSIST.

Next, this study will evaluate frustration as a dependent variable and examine the ways the treatment conditions impact the child's frustration levels. These conditions include target selection (complex targets vs. simple targets), the day of the week that treatment is conducted (e.g., Monday, Tuesday, Thursday, Friday), the session number within each day (e.g., morning [1, 2] and afternoon [3, 4] sessions), and the session number across the treatment period (early or later in the treatment timeline). Knowing the impact of specific conditions may be helpful in designing treatments to reduce frustration and stress. In addition, associations between baseline measures of CAS severity and average frustration levels during treatment will also be examined to determine whether CAS severity can predict frustration levels.

Finally, frustration will also be examined as an independent variable, specifically whether the average frustration level of a child during treatment will predict the speech accuracy gains they make. Frustration levels will be gathered and averaged for each child

and will be compared against the gains they made from baseline to post-treatment data. Whether children with CAS experience positive or negative impacts of their frustration may provide insight into how to manage children experiencing frustration in speech therapy based on its impact on performance and progress, as well as the overall effects of more intensive treatments for CAS.

To summarize, this study will address the following specific research questions:

1. What is the inter-rater reliability of clinician-determined frustration ratings based on participant behavior during treatment sessions?
2. To what extent is frustration level influenced by treatment factors such as treatment condition (simple versus complex targets), temporal variable such as day of week and time of day, and session number?
3. Is there a relationship between frustration level and CAS severity?
4. Does average frustration level during treatment predict response to treatment?

CHAPTER 2

METHODS

This study uses data collected as a part of a Phase 1 treatment study of ASSIST at Temple University in 2019. The data used for this study was collected in the first small-scale randomized controlled group treatment study which was focused on the parameter of complexity (simple vs. complex targets).

Participants

Seventeen participants were recruited based on the criteria listed in Tables 1, 2, and 3, and randomly assigned to one of four groups (immediate-simple, immediate-complex, delayed-simple, delayed-complex) (see Figure 1 for overview of the design of the treatment study). Participant data are provided in Table 4, organized by condition (simple vs. complex); the groups did not differ on any of the variables listed ($p > 0.25$, based on two-tailed independent samples t-tests). Each participant was involved in two weeks of direct treatment, which was conducted in 30-minute sessions, four times a day, four days a week in a summer camp setting. Outcome data was collected the week preceding treatment and the week following treatment (see T1, T2, and T3 in Figure 1).

Design

Figure 1 presents the design of the treatment study that forms the context of the present study. This study will use the frustration scores collected at the end of each treatment session by the treating clinician (four per day). These frustration scores will be

related to the treatment conditions (simple versus complex targets, session number, day of the week, and time of day) and the participant's CAS severity. Treatment targets were aimed to be within the optimal challenge levels in terms of target difficulty for each child. *Data Analysis* section below for details.

Inclusionary Criteria	
Age	4;0 to 9;11 at time of enrollment.
Primary language spoken at home	English (second language is not exclusionary).
Hearing status	Normal according to parent report; Pass audiometric hearing screen conducted by trained evaluator (response at 20dB in both ears at 1000, 2000, and 4000Hz).
Verbal output	At least 50 words with communicative intent according to parent report and SLP observation.
Speech sound disorder	<10th percentile on Goldman Fristoe Test of Articulation – 3rd Edition (GFTA-3).
Nonverbal cognition	Within 1.5 SD of mean on nonverbal subtests on the Reynold Intellectual Assessment Scales (RIAS).
Primary diagnosis	CAS (see Table 2)

Table 1. Inclusionary criteria for enrollment in ASSIST treatment study

Additional CAS-Specific Inclusionary Criteria	
Clinician judgement	As clinician judgement is the current gold standard for diagnosis, three independent expert SLPs will judge the child based on video recordings using a scale of 0-1-2 based on the perceptual characteristics of CAS. 0 means no CAS, 1 means possible CAS, and 2 means definite CAS. Inclusion requires an average score greater than 1.
Maximum performance tasks	An apraxia score of 1 or 2 on the maximum performance task (MPT) including sustained vowels and fricatives, and diadochokinetic speech tasks (Thoonen et al., 1996).

Table 2. Additional inclusionary criteria for enrollment in assist treatment study

Exclusionary Criteria	
Co-occurring diagnoses	Diagnosis of disorder that impacts communication or social interaction (e.g., autism).
Vision status	Vision impairment according to parental report.
Oral structure	Significant oral structure impairment as determined by evaluator with Oral Mechanism Examination (e.g., cleft palate).
Primary diagnosis	Dysarthria

Table 3. Exclusionary criteria for enrollment in ASSIST treatment study

Phase	Condition	wk1	wk2	wk3	wk4	wk5	wk6	wk7
Immediate (n=9)	Simple (n=4)	T1	8 hrs	8 hrs	T2			T3
	Complex (n=5)		8 hrs	8 hrs				
Delayed (n=8)	Simple (n=4)					8 hrs	8 hrs	
	Complex (n=4)					8 hrs	8 hrs	

Figure 1. Overview of design of the treatment study

Additionally, a randomized selection of 34 video-recorded sessions will be independently scored for frustration by the examiner to assess inter-rater reliability of the frustration rating scale used in the treatment study. The examiner will be blinded to the scores provided by the original analysts (treating clinicians).

Procedures

Frustration scores are recorded at the end of each session by the clinician, which occurs four times a day, four days a week, for a total of 32 scores per participant, assuming full attendance and no extenuating circumstances.

Treatment Procedures

ASSIST treatment sessions are separated into two parts: pre-practice and practice. Four targets are treated per 30-minute session from a larger set of 20, and targets can increase or decrease in difficulty based on performance, and new targets can replace old targets if they are determined to be mastered or too difficult. Mastered targets are also probed for maintenance at regular intervals and reintroduced if speech accuracy is lost. Both pre-practice and practice are operationalized step-by-step protocols.

Table 4. Participant information

child	sex	Age (y;m)	CAS dx	DEMSS	GFTA raw	GFTA SS	DEAP phonology SS	DEAP inconsistency	EVT SS	PPVT SS	RIAS NIX T	phase	condition
102	M	9;6	2.00	.	100	40	.	88%	20	64	32	1	simple
113	F	6;2	1.25	219	39	51	55	32%	91	99	41	1	simple
117	F	4;1	1.67	172	94	56	75	80%	88	90	44	1	simple
122	F	6;11	1.67	346	61	40	55	68%	74	79	43	1	simple
103	M	6;11	1.67	405	44	45	55	20%	76	104	64	2	simple
106	M	4;3	1.67	.	62	68	75	56%	128	127	89	2	simple
111	M	6;5	1.67	292	65	40	55	60%	87	103	52	2	simple
116	M	4;3	2.00	239	81	60	80	76%	88	83	43	2	simple
MEAN		***6;1***	***1.70***	***279***	***68***	***50***	***64***	***60%***	***82***	***94***	***51***		
stdev		***1;10***	***0.24***	***86***	***22***	***11***	***12***	***24%***	***30***	***19***	***18***		
104	M	6;0	1.33	305	69	40	55	72%	81	80	51	1	complex
105	M	9;1	2.00	343	33	40	55	48%	64	63	33	1	complex
114	M	8;7	2.00	288	24	44	55	52%	94	104	56	1	complex
115	M	5;0	1.25	315	75	51	65	52%	96	114	52	1	complex
121	M	4;5	2.00	78	97	48	75	76%	77	104	69	1	complex
101	M	6;6	2.00	353	44	53	60	52%	88	91	45	2	complex
107	M	5;0	2.00	219	75	51	70	76%	94	97	57	2	complex
119	M	7;2	2.00	187	93	40	55	52%	80	93	37	2	complex
120	M	6;2	2.00	229	102	40	55	72%	80	79	42	2	complex
MEAN		***6;5***	***1.84***	***257***	***68***	***45***	***61***	***61%***	***84***	***92***	***49***		
stdev		***1;7***	***0.31***	***88***	***28***	***6***	***8***	***12%***	***10***	***16***	***11***		

1. Pre-Practice

Pre-practice involves up to 15 minutes of establishing the ability to produce the target items. It uses a blocked practice system of up to five attempts per target before moving to the next target indicated on the scoresheet. Pre-practice uses integral stimulation ("watch me, listen to me, say what I say") and targets are produced in direct imitation with constant and regular prosody as well as immediate knowledge of results and knowledge of performance feedback. To move to the next stage (practice), all four targets must be produced accurately at least once in pre-practice, or the 15-minute time limit must elapse. It is possible to spend 15-minutes on a single target if the participant is unable to accurately produce it.

2. Practice

Practice is structured to elicit the targets in a way to facilitate learning with random practice of all four targets, variable prosody, and delayed and reduced feedback. It involves a hierarchal elicitation level system that depends on the participant's performance at the previous level, and includes direct imitation and constant prosody, direct imitation and varied prosody, delayed imitation and varied prosody, and finally independent production. These levels may go up or down a level during a session depending on the participant's performance.

Throughout practice, feedback is given in a structured manner and is systematically reduced throughout the session. Practice is provided in four 4-minute sets, with a one-minute break between the first two sets and the last two sets, and a two-minute break between the second and third sets.

There are several strategies used to mitigate potential frustration embedded in the treatment procedures. Target items are specifically chosen to be difficult, but within a range deemed appropriate based on the characteristics of the individual participant and organized in a list by complexity. Steps for target removal also exist in the case of a target being too difficult. Targets are presented in frames (short, functional phrases that include the target), which increase in complexity as the participant masters each frame, until the target is moved to maintenance. When a participant is unable to produce the target in the presented frame, the complexity level decreases to the lowest level, which is the target in isolation. If the participant is unable to produce the target in pre-practice as well as in the lowest elicitation level and the lowest frame level in practice, it is removed and replaced with another, less complex target before the next day. This is designed to reduce the frustration that may be caused by continued practice of a target that is not progressing. The treatment study is also held in a summer camp environment, where treatment sessions are scheduled between activities with other participants and camp counselors. This allows for time between intensive sessions to have fun and reduce potential frustration.

Data Collection

The treating clinician takes data in the moment and makes in-session, post-session, and end of day decisions regarding elicitation levels and target difficulty based on session data from practice.

Data for frustration is taken at the end of each session on a practice data sheet utilizing a 4-point scale (Table 5).

Frustration Rating Scale

The practice data collection sheet includes a frustration rating scale to be completed at the end of each session. The frustration rating scale is shown in Table 5 as written on the ASSIST data collection sheets. Noncompliance was provided as one way of judging frustration, but treating clinicians were expected to use reasonable clinical judgement based on their familiarity with the participant when providing the frustration score for that session.

Score	Scoring Criteria
0	No frustration, full compliance
1	Some frustration, occasional noncompliance
2	Significant frustration, frequent noncompliance
3	Marked frustration, noncompliant entire session

Table 5. ASSIST frustration rating scale

Data Analysis

To address Research Question 1 (RQ1) (What is the inter-rater reliability of clinician-determined frustration ratings based on participant behavior during treatment sessions?), we will examine intraclass correlations between the treating clinician and a second analysts based on 34 videos (2 per child, one from week 1, one from week 2). Specifically, we will use a single-rater two-way random-effects model for absolute agreement (Hallgren, 2012; Koo & Li, 2016). ICC values can range from -1 (perfect disagreement) to +1 (perfect agreement), with ICC = 0 indicating random (chance) agreement (Hallgren, 2012). Inter-rater reliability will be considered poor if ICC < 0.40,

fair if ICC < 0.59, good if ICC < 0.74, and excellent if ICC > 0.75 (Hallgren, 2012). If ICC is poor or fair, we will also examine percent agreement point-to-point and within one scale point.

To address Research Question 2 (RQ2) (To what extent is frustration level influenced by treatment factors such as treatment condition [simple versus complex targets], temporal variables such as day of week and time of day, and session number?) we will perform separate non-parametric analyses ($\alpha = 0.05$ for all analyses).

Specifically, to address the influence of treatment condition (a between-subject factor), we will perform a Kruskal-Wallis test on Wilcoxon rank sum scores for mean frustration scores across all sessions, with condition (Simple, Complex) as between-subjects factor. If targeting more complex utterances increases frustration, we expect significantly higher scores in the Complex condition than in the Simple condition.

To address the influence of day of the week, we will first perform the Cochran-Mantel-Haenszel chi-square test for comparisons involving more than two levels on mean frustration scores, with weekday (Monday, Tuesday, Thursday, Friday) as within-subjects factor. Wilcoxon paired signed rank tests will be used for pairwise comparisons. If frustration builds with repeated performance of difficult tasks (such as in ASSIST), then we expect significantly lower frustration scores for days following one or two days off (Monday, Thursday) than for second consecutive days (Tuesdays, Fridays). However, if frustration decreases with acclimation to difficult tasks over time, we may expect higher scores for days following breaks (Monday, Thursday), or for later weekdays (Thursday, Friday). If frustration builds over the course of the week, we expect higher frustration

scores for later weekdays (Thursday, Friday) than for earlier weekdays (Monday, Tuesday).

To address the influence of time of day, we will first perform the Cochran-Mantel-Haenszel chi-square test for comparisons involving more than two levels for mean frustration scores with the session number of the day (sessions 1, 2, 3, and 4) as within-subjects factor. Wilcoxon paired signed rank tests will be used for pairwise comparisons, including all session pairs as well as comparison between the mean of the morning sessions (sessions 1 and 2) and the mean of the afternoon sessions (sessions 3 and 4). If frustration increases with repeated performance of difficult tasks, we expect significantly higher frustration scores for afternoon sessions than for morning sessions. However, it is possible that acclimation to difficult tasks over time may reduce frustration, which would be represented by higher frustration scores in earlier sessions compared to later in the day. Additionally, if the inclusion of a mid-day break reduces frustration, we may expect lower mean frustration scores for session 3 compared to session 2 or 4.

Finally, to address the potential build-up of frustration over the course of the entire treatment period, we will use Spearman correlations to correlate session number (1-32) with mean frustration scores. If frustration increases over the intervention period, we expect a significant positive correlation (higher session numbers associated with higher mean frustration scores).

To address Research Question 3 (RQ3) (Is there a relationship between frustration level and CAS severity?), we will examine the Spearman correlation between mean

frustration level across all sessions and DEMSS score. If speech is more challenging for children with more severe CAS, then we expect a significant negative correlation between mean frustration scores and DEMSS scores (higher frustration scores associated with lower DEMSS scores, as lower DEMSS scores indicate more severe CAS).

To address Research Question 4 (RQ4) (Does average frustration level during treatment predict response to treatment?), we will examine the Spearman correlation between average frustration score and percent change from pre to post treatment. Treatment outcomes were determined using productions of treated and untreated target items (individualized for each participant based on motivational factors and optimal challenge level). If lower levels of frustration are related to better outcomes in treatment, we would expect to see a significant negative correlation.

CHAPTER 3

RESULTS

Frustration Rating Scale Inter-Rater Reliability

RQ1 was addressed with use of intraclass correlations (ICC) of frustration scores between the treating clinician and a second analyst, specifically a single-rater two-way random-effects model (Koo & Li, 2016), to determine inter-rater reliability of the frustration rating scale. The first week of treatment and the second week of treatment were examined separately and combined for all 17 participants, for a total of 34 treatment sessions analyzed. Additionally, frustration scores were examined with point-to-point agreement and agreement within one scale point to determine interrater score agreement.

Intraclass Correlation

Intraclass correlation measures for reliability include degrees of correlation (Table 6) and agreement between measures (Table 7). For frustration scores in the first treatment week across all participants, the ICC value was 0.185, indicative of poor reliability (see Data Analysis section for further discussion regarding ICC reliability range descriptors). For the second treatment week, the ICC value was 0.467, indicating fair reliability. Combined, both treatment weeks had an ICC value of 0.352, which is representative of poor reliability.

Treatment period	**ICC**	**95% CI**	**Reliability descriptor**
First half	0.185	[-0.305 – 0.600]	Poor
Second half	0.467	[-0.019 – 0.770]	Fair
Combined	0.352	[0.023 – 0.613]	Poor

Table 6. Frustration score ICC degrees of correlation

With regards to ICC agreement measures (Table 7), the first treatment week was determined to be not statistically significant, suggesting that the agreement between raters' frustration scores was not significantly different from chance. The second treatment week and the combined treatment weeks were determined to be statistically significant, indicating that the agreement between raters was significantly greater than chance.

Treatment period	**ICC agreement**	**Statistical significance**
First half	F(16,16.5) = 1.45, p = 0.229	Not significant
Second half	F(16,16.1) = 2.66, p = 0.0291	Significant
Combined	F(33,33.8) = 2.09, p = 0.0183	Significant

Table 7. Frustration score ICC agreement measures

Point Agreement

The frustration rating scores between raters were additionally examined through use of point-to-point agreement and agreement within one scale point (Table 8). To determine point-to-point agreements, the percentage of scores in exact agreement between raters was determined from the total (17 scores for first and second weeks, and 34 scores for the combined weeks). For agreement within one scale point, the percentage of scores within one scale difference of the second rater's score was determined. Point-to-point agreements ranged from 41% to 65%, while agreements within one scale point ranged from 88% to 94%.

	Point-to-point agreement		**Within one scale point**	
Treatment period	**Number**	**%**	**Number**	**%**
First half	7/17	41%	15/17	88%
Second half	11/17	65%	16/17	94%
Combined	18/34	53%	31/34	91%

Table 8. Frustration scale point agreements

Treatment Factors and Frustration Levels

RQ2 was addressed with separate non-parametric analyses (Kruskal-Wallis, Cochran-Mantel-Haenszel, and Wilcoxon paired signed ranks tests) for comparisons of treatment conditions and temporal variables, and Spearman correlations for assessing the association between frustration and session number.

Simple vs. Complex Targets

The influence of treatment condition was analyzed with a Kruskal-Wallis one-way ANOVA on Wilcoxon rank sum scores of mean frustration levels across all sessions with the condition of target complexity (simple vs. complex) as the between-subjects factor (Table 9). The non-parametric analysis revealed no significant difference for average frustration scores between the complex condition and the simple condition (χ^2= 0.189, p = 0.664). No effect of complexity was revealed on frustration.

Condition	**n**	**mean**	**SD**	**median**	**min**	**max**
Complex	9	0.382	0.475	0.355	0	1.471
Simple	8	0.488	0.467	0.344	0	1.310

Table 9. Descriptive data for frustration scores by treatment condition

Day of the Week

The influence of the first temporal condition (day of the week) was first analyzed with a Cochran-Mantel-Haenszel chi-square test for variables with more than two levels, with treatment day (Monday, Tuesday, Thursday, Friday) as the within-subjects factor (Table 10). The non-parametric analysis revealed no significant difference (χ^2= 2.781, p = 0.427). Analysis of all pair-wise comparisons using Wilcoxon paired signed rank tests confirmed that there were no significant differences between days (all p-values > 0.30).

Further, we also compared the first half (average of Monday and Tuesday) vs. second half of the week (average of Thursday and Friday) using the Wilcoxon paired signed rank test. There was no significant difference for mean frustration scores between the first half of the week vs. the second half of the week (Wilcoxon S = 15, p = 0.366). Finally, no differences were observed when comparing the first day after a break (average of Monday and Thursday) and the second day after a break (average of Tuesday and Friday) (Wilcoxon S = 1, p = 0.964).

Thus, no effect of day of week was revealed on frustration.

Condition	**N**	**mean**	**SD**	**median**	**min**	**max**
Monday	17	0.480	0.503	0.375	0	1.750
Tuesday	17	0.483	0.556	0.250	0	1.714
Thursday	17	0.413	0.501	0.250	0	1.500
Friday	17	0.377	0.435	0.167	0	1.250

Table 10. Descriptive data for frustration levels by day of the week

Time of Day

The influence of the second temporal condition (time of day) was first analyzed with a Cochran-Mantel-Haenszel chi-square test for variables with more than two levels, with session number within the day of treatment (sessions 1, 2, 3, and 4) as the within-subjects factor (Table 11). The non-parametric analysis revealed a non-significant trend toward differences for average frustration scores between sessions ($\chi 2 = 7.210$, p = 0.066). Follow-up inspection of all pair-wise comparisons using Wilcoxon paired signed rank tests (see Table 12 below): Session 1 had significantly lower frustration scores than Session 4 and marginally lower scores than Session 3; Session 2 also had marginally lower frustration scores than Session 4. This pattern suggests a possible effect of time of

day. To follow up on this effect, we compared the mean of the morning sessions (1 and 2) to the mean of the afternoon sessions (3 and 4) using the Wilcoxon paired signed rank test, which revealed significantly higher frustration scores in the afternoon than in the morning.

Thus, frustration levels appear to increase over the course of the day.

Condition	**n**	**mean**	**SD**	**median**	**min**	**max**
Session 1	17	0.371	0.417	0.250	0	1.400
Session 2	17	0.382	0.476	0.143	0	1.500
Session 3	17	0.456	0.477	0.333	0	1.429
Session 4	17	0.509	0.543	0.375	0	1.667

Table 11. Descriptive data for frustration scores by daily session

	n	**Difference (SD)**	**Wilcoxon S**	**p-value**
Session 1 vs. 2	17	-0.012 (0.220)	-6.5	0.593
Session 1 vs. 3	17	-0.085 (0.145)	-20.5	0.066
Session 1 vs. 4	17	-0.139 (0.229)	-32.5	**0.020**
Session 2 vs. 3	17	-0.073 (0.247)	-13.5	0.249
Session 2 vs. 4	17	-0.127 (0.279)	-27.5	0.055
Session 3 vs. 4	17	-0.054 (0.180)	-13.5	0.250
Morning vs. Afternoon	17	-0.106 (0.181)	-35.5	**0.0245**

Table 12. Pair-wise differences and comparisons for session effects on frustration level

Session Number

To analyze the potential progressive culmination of frustration over the course of the treatment period, a Spearman rank correlation between group mean frustration scores and session number (1-32) was completed. This correlation revealed a significant

negative relationship (Spearman ρ = -0.362, p = 0.042). Across all children, frustration level decreased over time (Figure 2).

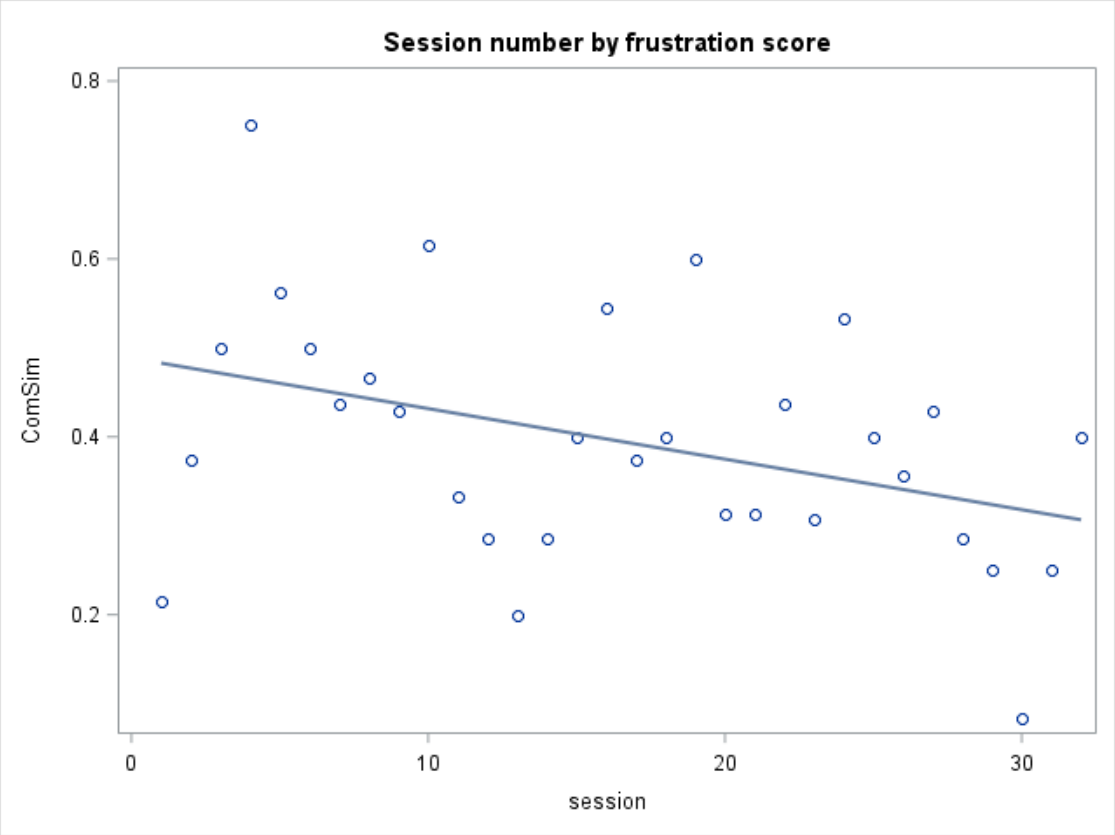

Figure 2. Session number by frustration score*

*Please note the y-axis indicates mean frustration levels across both treatment conditions (ComSim). The x-axis indicates session number.

CAS Severity and Frustration Levels

RQ3 was addressed a Spearman rank correlation between mean frustration scores and DEMSS scores for participants with DEMSS scores (Table 13). Lower DEMSS scores indicate more severe CAS. Of the 17 total participants, the DEMSS was completed for 15 prior to the start of treatment; the correlation for RQ3 was completed using data

for those 15 participants. The correlation revealed no relationship between CAS severity and frustration during sessions (Spearman $\rho = 0.066$, $p = 0.814$).

Participant number	DEMSS score	Mean frustration
101	353	1.471
103	405	0.000
104	305	0.667
105	343	0.000
107	219	0.000
111	292	0.667
113	219	0.174
114	288	0.033
115	315	0.435
116	239	0.344
117	172	1.310
119	187	0.043
120	229	0.355
121	78	0.433
122	346	1.000

Table 13. Participant DEMSS and mean frustration scores

Frustration Levels and Treatment Outcomes

RQ4 was addressed with a Spearman rank correlation between mean frustration scores and the percent change for all treated and untreated items (post-treatment – pre-treatment) for the 15 participants with complete data (Table 14). The correlation revealed no significant relationship for all participants (Spearman $\rho = -0.249$, $p = 0.371$), nor when considering only items that were practiced in treatment, rather than all untreated and treated items (Spearman $\rho = -0.267$, $p = 0.337$).

Item set	Condition	N	ρ	p-value
ALL	Both	15	-0.249	0.371
	Simple	6	0.439	0.384
	Complex	9	**-0.727**	**0.027**
TREATED	Both	15	-0.267	0.337
	Simple	6	0.663	0.151
	Complex	9	**-0.880**	**0.002**

Table 14. Correlations between frustration levels and treatment progress across treatment conditions and item sets

An additional analysis of each treatment condition (simple vs. complex) was completed with a Spearman rank correlation between mean frustration scores separated by condition and percent change for all items (Table 14). When analyzed separately for each treatment condition, there was a significant negative correlation for the complex condition for mean frustration score for all items (Figure 3) (Spearman $\rho = -0.727$, $p = 0.027$), as well as for treated items (Figure 4) (Spearman $\rho = -0.880$, $p = 0.002$). No significant correlation was observed in the simple condition for mean frustration score for all items (Spearman $\rho = 0.439$, $p = 0.384$) nor for treated items only (Spearman $\rho = -0.663$, $p = 0.151$). There was a greater positive change in treatment outcomes associated with lower frustration scores among participants in the complex treatment target condition. In the simple condition and the combined conditions, there was no association between frustration levels and treatment progress.

More detailed frustration data organized by participant and various conditions are provided in Table 15.

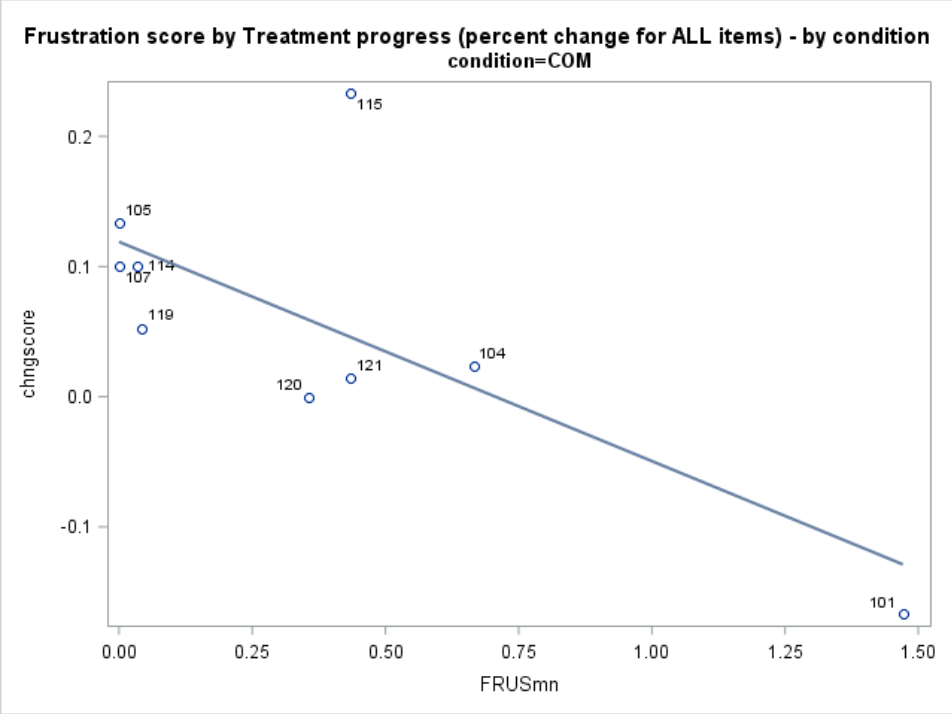

Figure 3. Mean frustration score by treatment progress in the complex target condition for treated and untreated items

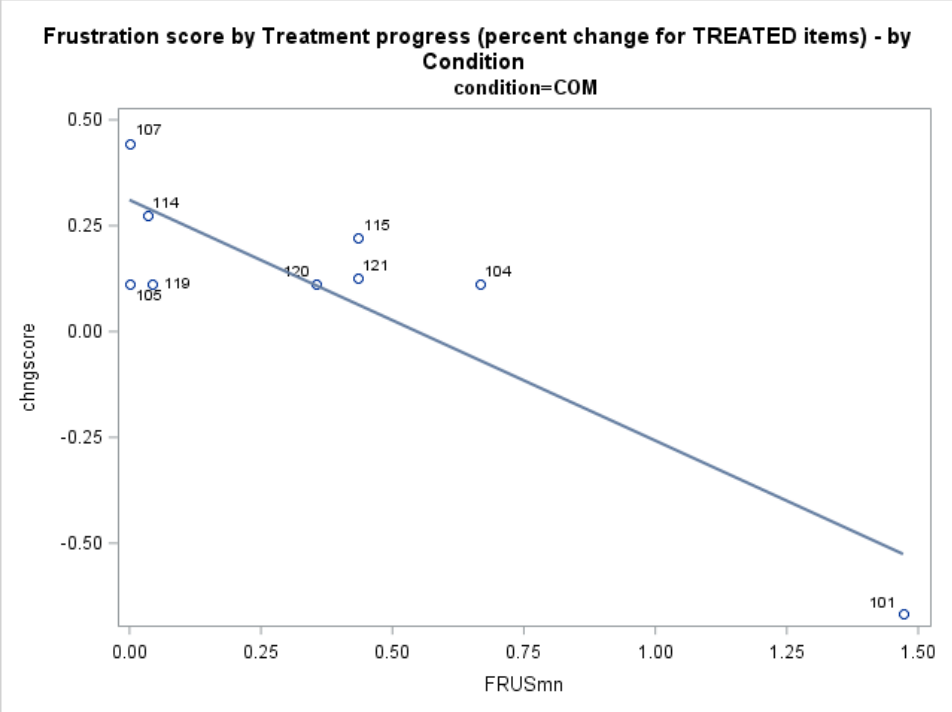

Figure 4. Mean frustration score by treatment progress in the complex target condition for treated items only

Table 15. Frustration data by participant

child	FRUS	FRUS sd	# of 2 scores	# of 3 scores	# scores ≥ 2	# sessions	% sessions 3	% sessions ≥ 2	Monday	Tuesday	Thursday	Friday	s1	s2	s3	s4
101	1.471	0.624	6	1	7	17	0.0588	0.4118	1.750	1.714	1.000	1.000	1.400	1.500	1.400	1.667
102	0.063	0.246	0	0	0	32	0.0000	0.0000	0.000	0.250	0.000	0.000	0.000	0.000	0.125	0.125
103	0.000	0.000	0	0	0	28	0.0000	0.0000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
104	0.667	0.758	5	0	5	30	0.0000	0.1667	1.000	0.286	0.875	0.500	0.500	0.571	0.875	0.714
105	0.000	0.000	0	0	0	29	0.0000	0.0000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
106	0.344	0.545	1	0	1	32	0.0000	0.0313	0.875	0.125	0.375	0.000	0.250	0.500	0.250	0.375
107	0.000	0.000	0	0	0	30	0.0000	0.0000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
111	0.667	0.816	3	0	3	15	0.0000	0.2000	0.250	1.333	1.000	1.000	0.600	0.000	0.750	1.000
113	0.174	0.388	0	0	0	23	0.0000	0.0000	0.250	0.250	0.000	0.143	0.200	0.143	0.333	0.000
114	0.033	0.183	0	0	0	30	0.0000	0.0000	0.143	0.000	0.000	0.000	0.143	0.000	0.000	0.000
115	0.435	0.507	0	0	0	23	0.0000	0.0000	0.375	0.500	0.500	0.333	0.500	0.500	0.333	0.400
116	0.344	0.483	0	0	0	32	0.0000	0.0000	0.375	0.125	0.250	0.625	0.375	0.125	0.500	0.375
117	1.310	0.471	9	0	9	29	0.0000	0.3103	1.143	1.375	1.500	1.250	1.125	1.143	1.429	1.571
119	0.043	0.209	0	0	0	23	0.0000	0.0000	0.000	0.000	0.000	0.250	0.000	0.000	0.000	0.167
120	0.355	0.608	2	0	2	31	0.0000	0.0645	0.500	0.750	0.000	0.143	0.250	0.375	0.375	0.429
121	0.433	0.774	2	1	3	30	0.0333	0.1000	0.625	0.500	0.375	0.167	0.125	0.500	0.375	0.833
122	1.000	0.277	1	0	1	27	0.0000	0.0370	0.875	1.000	1.143	1.000	0.833	1.143	1.000	1.000

CHAPTER 4

DISCUSSION

The primary goals of the current study were to 1) determine the interrater reliability of frustration rating scale used in the ASSIST treatment study, 2) evaluate the influence of various treatment factors on the frustration levels of the participants, 3) assess the relationship between CAS severity and frustration level in ASSIST, and 4) examine the relationship between frustration level and treatment progress following ASSIST.

RQ1: Reliability of ASSIST Frustration Rating Scale

During the ASSIST treatment study, frustration levels were taken for each participant by the treating clinician following each treatment session using the 4-point scale discussed in the *Methods* section above. Due to its subjective nature, the reliability of the scale was necessary to be evaluated as it is utilized to monitor participant wellbeing throughout the study.

Intraclass Correlation of Frustration Scores

Frustration levels were independently assessed by a second rater blinded to the original scores for 34 randomly selected treatment sessions (one session per week randomly selected for each participant). The two score sets were compared using intraclass correlations (ICC), specifically a single-rater two-way random effects model. It was determined that frustration scores given in the first week of treatment had poor reliability and agreement was not significantly different from chance. The second week

had fair reliability, with agreement significantly greater than chance. When both weeks were combined, the reliability was deemed as poor, but the agreement was significantly greater than chance.

The results of the ICC analysis suggest that the frustration rating scale used during ASSIST has a poor to fair level of reliability. This could be related to several reasons, including both scale-specific factors and clinician-specific factors.

Possible scale-specific factors include the fact that the conditions provided for each frustration point may have been too vague or nonspecific, which opens the door for interpretations which may not have been shared among treating clinicians. "Noncompliance" as the example descriptor for frustration in the rating scale may have also influenced clinicians to not consider possible frustration-indicative behaviors unrelated to compliance (e.g., impatience or irritation with continued participation) in their score. It also may have influenced clinicians to consider noncompliance unrelated to frustration (e.g., noncompliance due to boredom or playful mood) in their score.

Possible clinician-specific factors include different perceptions of frustration in treatment, which could be related to clinician inexperience (all treating clinicians were graduate students with less than two years of clinical experience) or general discrepancies in perceptions of frustration – what frustration is, as well as the difference between each scale point. Treating clinicians could have also had some recency bias when scoring frustration since scores were determined at the end of the session. As a result, behaviors that occurred in the earlier part of the session may be weighted less strongly than behaviors that occurred later in the session. Since the second rater judged the frustration based on video recordings, frustration could be tracked more easily through the session.

Treating clinicians also worked with the same participants across the treatment period, which could have influenced their frustration score decisions. If a participant was exhibiting highly frustrated behaviors in prior sessions, the clinician may have been primed to observe certain behaviors that the second rater was unaware of. Treating clinicians were also more knowledgeable of the child's temperament and may be more adept at discerning specific behaviors. The second rater had no prior experience with the participants, and therefore less capable of determining if behaviors were baseline or frustration-specific.

Additionally, treating clinicians were responsible for providing the treatment, and therefore tasked with other focuses during the session including presenting targets, cueing for accuracy, data collection, and other treatment-related duties. It is possible for the treating clinician to miss or misjudge behaviors due to their need to multitask through the session, while the second rater was focused only on observing frustration behaviors. It is also possible for clinicians to have forgotten to score frustration immediately after some sessions, forcing them to provide a score later in the day, which would reduce the accuracy of the score.

The ICC agreement measures indicated that the first week treatment resulted in frustration scores that were not significantly different from chance. This could be related to factors such as clinician comfort and competence with the treatment protocols and the frustration rating scale, as well as familiarity of the specific traits of the participants they are judging. It is possible that the treating clinicians became more accurate in their frustration scoring with greater experience providing the treatment to the participants, resulting in better agreement in the second week.

Point Agreement Measures

Regarding the point agreement measures, it was found that point-to-point agreements between the original rater scores and the second rater scores were low but substantially higher and acceptable when considering agreements within one scale point. The agreements within one scale point ranged from 88% to 94%, while the point-to-point agreements did not surpass 65%. This suggests that while two scorers may not reach the exact same conclusions for frustration levels, they typically provide scores that are within a single scale point of each other. Due to the subjective nature of the rating scale, and the impact of clinician-specific factors to their scores, some discrepancy in scoring is likely unavoidable. Greater concern for the efficacy of the scale would arise if two independent scorers frequently provided scores that had a discrepancy of two or more points.

Considerations and Implications

While the reliability of the frustration rating scale has been determined by the analyses to be poor to fair, it is still a useful measure of frustration given the constraints required by the scale such as feasibility of use and the acceptable agreement within one scale point. The acceptable agreement within one scale point supports consideration of the additional research questions addressed below.

Some level of difference is expected due to the subjective nature of the scale, but elements could be addressed to further reduce discrepancy. By removing noncompliance from the frustration scale descriptions, and possibly creating a separate scale to judge noncompliance exclusively, could reduce interference of non-frustration related behaviors in the judgement of frustration. More extensive training on frustration (such as different possible presentations of frustration levels and types) could increase clinician

ability to recognize behaviors more consistently. Additionally, increasing familiarity with treatment protocols (through additional practice) and participant temperaments (through meeting participants before treatment, or watching videos of their pre-treatment evaluations) prior to the start of treatment may also reduce inaccurate scoring.

RQ2: Influence of Treatment Factors on Frustration

Various treatment factors during ASSIST were evaluated based on their influence on the mean frustration levels across all participants in the study. Assessing the effects of treatment factors on potential exacerbation or reduction of frustration is important to determine optimal treatment conditions in the clinical setting. Ultimately, ASSIST is designed for future clinical usage, and understanding which, if any, controllable variables are related to greater emotional wellbeing for the individual receiving the treatment is important for client-centered practice.

Conditions analyzed include treatment condition (simple vs. complex targets), day of the week (Monday, Tuesday, Thursday, Friday), time of day (sessions 1 and 2 in the morning, sessions 3 and 4 in the afternoon), and session number (across two-week treatment timeline). Complexity and day of the week did not have an effect on frustration level among the participants. This suggests that greater complexity does not increase frustration levels, and that frustration neither builds nor reduces over the course of the week.

Time of Day

When examining time of day, we observed a trend towards greater frustration later in the day. Further inspection revealed that afternoon sessions, and especially

session 4, resulted in higher frustration levels than morning sessions. This increase may reflect a build-up of frustration from repeated performance of difficult tasks, rather than an acclimation to these difficult tasks. An alternative interpretation is that this effect is related to time of day itself and children's circadian rhythms, rather than (or in addition to) how many sessions have been completed that day. Perhaps children are more resilient in the morning than in the afternoon. This possibility is consistent with observations that cortisol hormone levels (a stress-response hormone) show a typical diurnal pattern of higher levels in the morning and a subsequent tapering across the day (e.g., Cicchetti & Rogosch, 2001; Klimes-Dougan et al., 2001). Although the relationship between stress/resilience and cortisol levels is complex (e.g., Rao & Androulakis, 2019), if higher levels of cortisol are associated with preparation to face challenges (e.g., Laures-Gore et al., 2019; Schlotz et a., 2004) and thus perhaps greater resilience in the face of stressors, then it is possible that the observed time-of-day effect reflects, at least in part, these circadian endocrinological variations.

Session Number

Treatment occurred four times per day, four days a week for two weeks, resulting in 32 total sessions. Frustration level across all participants was found to decrease over time during treatment. This suggests that frustration does not necessarily progressively cumulate over time when participating in difficult tasks such as intensive treatment. It is possible that acclimation to the routine and structure of the treatment, as well as increased comfort with the treating clinician, aided in the reduction of frustration over time. Another possibility is the treating clinicians developed more successful behavior

management strategies to prevent frustration as they became more familiar with the participants.

Considerations and Implications

It is important to consider that the level of treatment intensity in the ASSIST treatment study was considerably greater than what is present in typical CAS treatments, including other intensive summer camps. Many other summer camps exist for CAS treatment but due to the rigorous schedule and rigid treatment protocolization of the ASSIST treatment study, the data used in the current study is derived from an exceptionally intensive methodology.

Regarding target complexity and the weekday of treatment, it is suggested that these factors may not play a role in the exacerbation of frustration among the participants. While frustration did occur during some sessions, the triggers for such heightened emotions and subsequent behaviors may be more related to other factors. The drill-based format of the treatment included very strict protocols, including limited break times and a pre-set feedback schedule during treatment. Children that experience heightened sensitivity, anxiety, or desire for perfectionism at baseline may also experience greater frustration in treatment due to these factors.

Afternoon sessions resulted in higher frustration levels than morning sessions, suggesting that scheduling sessions in the morning may be preferred to reduce frustration, if possible, in the clinical context. It is also important to note that frustration did not increase with two sessions per day (i.e., two morning sessions), suggesting that treatment intensity can be increased from singular daily sessions without frustration effects.

Frustration levels were found to decrease over time among all participants, suggesting that consecutive involvement in difficult tasks may result in acclimation to the rigor rather than a progressive build-up of frustration. It is likely that the start of a new routine consisting of frequent and regular intensive treatment sessions in an unfamiliar location led to heightened frustration in the beginning, which tapered downwards as the participants became more familiar with the setting and expectations. When considering clinical practice, the downward trend of frustration with time suggests that children with CAS may initially react to treatment with behaviors such as noncompliance or tearfulness. This early reaction should not necessarily dissuade the clinician or the parent from continuing sessions as the child will likely become less frustrated as treatment progresses. If significant frustration does not decrease with time, other contributing factors should be considered.

RQ3: Effects of CAS Severity on Frustration

The DEMSS was administered to all participants prior to treatment start to determine eligibility for the study as well as CAS severity. Considering the increased communicative difficulties that children with more severe CAS may experience, it was hypothesized that increased severity may be related to higher frustration levels. However, the analysis revealed that there was no relationship between CAS severity and frustration levels among participants, suggesting that increased levels of communicative difficulties may not necessarily heighten feelings of anger and annoyance when participating in intensive speech therapy.

Considerations and Implications

As discussed briefly in the *Methods* section, the treatment targets were developed in terms of difficulty with consideration of the optimal challenge level for each child. A target that may be exceedingly challenging for one participant may be too simple for another. Due to the optimization of target difficulties for each child, the relationship between CAS severity and frustration may not be adequately assessed with these data as target difficulty was designed to be relatively equal for each child with consideration of their CAS severity. Therefore, while no effect was revealed in the present study, the results may not be indicative of the true relationship between CAS severity and frustration in treatment. Instead, the lack of an effect may suggest that the target selection process did in fact result in creation of sets at the optimal challenge level for each child.

RQ4: Relationship Between Frustration and Treatment Progress

Frustration levels were found to only be associated with a change in treatment outcomes for participants in the complex condition for both treated and untreated items. Targets in the complex condition were embedded in more complex carrier phrases than those in the simple condition, and thus this condition was designed to be more challenging than the simple condition. The results indicate that within the participants assigned to the complex condition, children who exhibited less frustration had better treatment outcomes compared to those with higher mean levels of frustration. It is possible that participants who had fewer instances of frustrated behavior ultimately participated in treatment more, contributing to better outcomes due to increased practice. One could also surmise that participants with higher frustration levels may have

associated treatment with negative feelings. This could potentially result in the conscious or subconscious decision to not use skills practiced in the sessions outside of treatment, hindering the maintenance and generalization of aforementioned skills. However, since no relationship between frustration and treatment progress was observed in the simple condition, these possibilities may not necessarily be the case, or at least not unless a certain level of frustration is reached (due to more challenging utterances being practiced).

Another possible explanation for better outcomes among these participants is related to the disposition of the child when presented with additional challenges. In discussion of grit in the *Introduction* section above, it was noted that greater success rates are related to the ability to pursue goals despite obstacles (such as more complex targets) with the suppression of behaviors (such as frustration) that would deter from accomplishing said goal (Duckworth & Gross, 2014). Perhaps "grittier" children were more capable of managing the added challenge of complex utterances due to their motivation to succeed, resulting in their ability to suppress frustration-related behaviors in pursuit of that goal. These children may have experienced feelings of frustration when faced with obstacles but did not display related behaviors, resulting in lower frustration scores. Based on that assumption, children with a different disposition may not have had the same drive to succeed in treatment and therefore did not suppress behaviors related to frustration when faced with the additional challenge of complex utterances, which would result in higher clinician-rated frustration levels.

To better visualize the difference between "grittier" children and the other participants in the complex condition, consider the children as a cup that can hold a

certain threshold of water, which represents the burden of the challenges they are experiencing, such as intensive treatment. Perhaps "grittier" children have a slightly larger cup, and are able to hold additional water (e.g., complex utterances) compared to other children, who would overflow, resulting in frustration-related behaviors as they are unable to manage the added burden. This would explain why this phenomenon is not observed in the simple condition, as the water representing the intensive treatment does fill up their cup to some extent, but within the threshold they are all able to manage so differences in outcomes are not observed.

Disposition, if we are to assume this possibility as a likely explanation, is suggested to be a contributing outside variable influencing both treatment outcomes and frustration levels among participants in the complex condition. When given the additional challenge of complex utterances, children who were more focused on pursuing success in treatment (i.e., "grittier" participants) may have had better outcomes as well as reduced frustration-related behaviors. Under this assumption, there may not be a direct relationship between frustration level and treatment outcomes, which is supported by the results of the simple and combined condition correlations, as the difference exists only in the complex condition.

Considerations and Implications

Without further study into the effects of frustration levels in children with CAS on treatment outcomes, firm conclusions cannot be drawn regarding their relationship. Several other factors could impact frustration levels or treatment outcomes, and some may influence both concurrently. Child disposition is simply one possibility of many potential contributing factors. Further exploration is warranted to determine those factors

to optimize outcomes for clinical practice. While the direct relationship to treatment outcomes is unclear at this time, it is still worth considering potential frustration levels in clinical decisions for the sake of client wellbeing alone.

General Implications

Although its causes and implications in relation to speech and language therapy have not been adequately researched, frustration is often used as a subjective descriptor in the field's clinical documentation for evaluation and treatment. It acts as a "catch-all" term to address unagreeable behavior, agitation, lack of motivation, and low participation in both child and adult clients, which may then be used to justify future clinical decisions, including cessation of services if clients present with recurrent frustration.

The present study revealed the relatively poor reliability of a simple frustration rating scale used in an intensive treatment study for children with CAS, suggesting that perceptions of frustration vary between clinicians. It also suggests that defining frustration by specific behaviors (e.g., noncompliance) may be too restrictive given the variability in expression of frustration. Additionally, the relationships between severity of communication disorder, frustration-related behaviors, treatment conditions, time spent in treatment, and treatment outcomes are unclear due to the poor operational definition of frustration. Without a more standard clinical definition of frustration and a better understanding of how frustration impacts and is impacted by treatment, the use of the term in clinical practice is not as valuable as it may seem. Descriptors of behaviors (e.g., noncompliant, distracted, tearful) rather than perceived emotions (e.g., frustrated, sad, stressed, angry) may more accurately describe a clinical scenario without assuming

underlying emotional basis. For the sake of best clinical practice, we as a field must carefully consider our verbiage to ensure accurate representations of our clients without assumptions or imprecise conclusions.

Limitations

Several limitations were present in the current study and should be taken into consideration with the results. First, the implications of the poor reliability of the frustration rating scale addressed with RQ1 on the other three research questions cannot be understated. Since the reliability of the scale utilized to determine the scores that were subsequently used in other analyses, the results of those analyses may not accurately reflect the relationships studied. Therefore, conclusions drawn from the results of the current study should be considered tentative.

A small sample size was used for the data analyzed in the present study. A total of 17 participants were involved in the ASSIST treatment study, and only 15 had full sets of data. When analyzing by treatment condition, the sample size is cut roughly in half. CAS is a rare speech disorder and as a result, it may be difficult to conduct a treatment study with a larger group, although the present sample is relatively large in the context of the CAS treatment literature, where an average sample size of two children is not uncommon (e.g., Murray et al., 2014), and comparable to sample sizes in the CAS diagnostic literature, where a median sample size of 11 is common (Murray et al., 2021). To account for the small sample sizes and likely violations of assumptions underlying parametric tests, we used non-parametric statistics. Nevertheless, the small sample sizes may have contributed to the obtained null effects. More generally, a small sample size is unlikely to

accurately represent the population, and therefore even the observed significant effects should be viewed with caution and require replication with a larger sample.

Additionally, the level of intensity in this particular treatment format is far higher than typically seen in CAS treatments, let alone typical speech therapy. While the present study sought to address frustration in the setting of an intensive treatment, the results may not be as applicable to individuals participating in treatments of lower intensity. Without frustration data from treatment sessions more representative of the norm, one may not be able to make direct comparisons between participants in the ASSIST treatment study and typical individuals receiving speech therapy.

Directions for Future Research

The topic of frustration as it relates to speech and language treatment includes many opportunities for future research. Development of a more reliable frustration scoring system that remains feasible in clinical practice would be beneficial in standardizing the clinical definition of frustration and reducing inconsistency among clinicians. To develop such a scale, frustration must be addressed as the primary focus of a study. The present study used frustration data from research primarily focused on other aspects of treatment, and the data for frustration was likely limited as a result.

A study dedicated to frustration in speech therapy could include more extensive measurements of child baseline temperament and personality, through parent questionnaires or personality assessments, to address the possible factor of disposition in frustration and treatment outcomes. Determining baselines for frustration-related

behaviors could also be beneficial in such a study to ensure that behaviors in treatment being judged as frustration are truly representative of frustration for that child. The addition of bioindicators of frustration, such as heart rate monitors or skin conductance would provide an additional measurement of frustration beyond clinician judgement and would enable a more precise view of the relationship between particular events and frustration, given the continuous nature of such biophysiological measures.

Involving child self-assessment of frustration and/or emotion before and after treatment sessions could reveal internal elements related to behaviors in treatment. These self-ratings could also call greater attention to the child's own frustration-related behaviors exhibited in treatment. Research dedicated to the impact of awareness of frustration during treatment could reveal possible connections between self-awareness, control of outward behavioral expression of frustration, and treatment.

Abandonment of practice due to frustration intolerance (i.e., how long it takes for a child to display significant noncompliance or other behaviors incompatible with treatment when internally experiencing a buildup of frustration) could be another area of study. Rather than focusing solely on perceived level of frustration across an entire treatment session, it may be beneficial to view frustration from a temporal standpoint (e.g., time between first outward expression of frustration to significantly unagreeable behavior, short outbursts of substantial frustration compared to long periods of lower frustration levels, etc.) to examine potential effects on treatment.

Within the scope of the ASSIST treatment study, the following treatment period was conducted virtually via Zoom in 2021 during the COVID-19 pandemic. This phase

of the study used different treatment conditions (massed [as in the present study] vs. distributed practice) than the conditions in the current study, but all other aspects of the treatment protocols remained the same given the constraints of teletherapy. A comparison of frustration levels between in-person and virtual treatments would be of interest given the recent increase of virtual speech and language treatment.

Additionally, a study of ASSIST with a larger number of treatment sessions per clinician (through a longer treatment period and/or a larger participant pool) may provide greater frustration data per clinician. Inter-rater reliability testing per rater rather than with all raters combined may yield different results. The present study used a single second analyst, but five original scorers. Therefore, the second analyst's frustration scores are judged by a single interpretation of frustration but was compared to five potentially different perceptions of frustration. By comparing a second analyst's scores to a single rater rather than five separate clinicians, further information regarding reliability may be revealed without the potential confounding factor of inconsistencies between clinicians.

Further research into frustration and its relationship with treatment for speech, language, fluency, cognitive communication, swallowing, and other areas of focus among children and adults is warranted and recommended for increased breadth of knowledge in clinical practice and to add to the limited current body of research.

CHAPTER 5

CONCLUSION

Frustration is a term often used but minimally researched in speech-language pathology. This is the first study to systematically examine frustration in relation to treatment for children with CAS. The study first sought to examine the interrater reliability of a clinician-rated frustration scoring system used in intensive treatment for children with CAS. It was revealed that the frustration rating scale used may not be reliable between raters. Additionally, frustration levels were analyzed to determine potential relationships with treatment conditions, temporal conditions, session number, CAS severity, and treatment outcomes. This study revealed possible relationships between session number and frustration, between time-of-day and frustration, as well as frustration level and treatment outcomes in complex treatment conditions. However, further research into frustration is warranted due to limitations of the present study and the scarce body of research on the topic.

REFERENCES CITED

Adani, S., & Cepanec, M. (2019). Sex differences in early communication development: behavioral and neurobiological indicators of more vulnerable communication system development in boys. *Croatian Medical Journal, 60*(2), 141–149. doi:10.3325/cmj.2019.60.141

American Speech-Language-Hearing Association. (2007). *Childhood apraxia of speech* [Technical Report]. Available from www.asha.org/policy/.

American Speech-Language-Hearing Association. (2020). *Schools Survey Report: SLP Caseload and Workload Characteristics Report.* Available at www.asha.org.

Apraxia Kids. (2021). *School-Based Training.* https://www.apraxia-kids.org/virtual-education-series/.

Ballard, K. J., Robin, D. A., McCabe, P. J., & McDonald, J. (2010). A treatment for dysprosody in childhood apraxia of speech. *Journal of Speech, Language, and Hearing Research, 53*(3), 1227-1245. doi:10.1044/1092-4388(2010/09-0130)

Cicchetti, D., & Rogosch, F. A. (2001). The impact of child maltreatment and psychopathology on neuroendocrine functioning. *Development and Psychopathology, 13* (4), 783-804.

Davis, B., & Velleman, S. L. (2000). Differential diagnosis and treatment of developmental apraxia of speech in infants and toddlers. Infant-Toddler Intervention: The Transdisciplinary Journal, 10, 177–192.

Duckworth, A., & Gross. J. J. (2014). Self-control and grit: Related but separable determinants of success. *Current Directions in Psychological Science, 23*(5), 319-325. doi:10.1177/0963721414541462

Gildersleeve-Neumann, C. (2007). Treatment of childhood apraxia of speech: A description of integral stimulation and motor learning. *ASHA Leader, 12*(15), 10-30. doi:10.1044/leader.FTR3.12152007.10

Hall, P. K., Jordan, L. S., & Robin, D. A. (1993). *Developmental apraxia of speech: Theory and clinical practice.* Pro-Ed.

Hallgren, K. A., (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology, 8*, 23-34.

Hokanson, J. E., & Burgess, M. (1964). Effects of physiological arousal level, frustration, and task complexity of performance. *Journal of Abnormal and Social Psychology, 68*(6), 698-702.

Klimes–Dougan, B., Hastings, P. D., Granger, D. A., Usher, B. A., & Zahn–Waxler, C. (2001). Adrenocortical activity in at-risk and normally developing adolescents: Individual differences in salivary cortisol basal levels, diurnal variation, and responses to social challenges. *Development and Psychopathology, 13* (3), 695-719.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15* (2), 155-163.

Laures-Gore, J., Cahana-Amitay, D., & Buchanan, T. W. (2019). Diurnal cortisol dynamics, perceived stress, and language production in aphasia. *Journal of Speech, Language, and Hearing Research, 62* (5), 1416-1426.

Lewis, B. A., Freebairn, L. A., Hansen, A. J., Iyengar, S. K., & Taylor, H. G. (2004). School-age follow-up of children with childhood apraxia of speech. *Language, Speech, and Hearing Services in Schools, 35*(2), 122-140. doi:10.1044/0161-1461(2004/014)

Longo, Y., Gunz, A., Curtis, G. J., & Farsides, T. (2016). Measuring need satisfaction and frustration in educational and work contexts: The Need Satisfaction and Frustration Scale (NSFS). *Journal of Happiness Studies: An Interdisciplinary Forum on Subjective Well-Being, 17*(1), 295–317. doi:10.1007/s10902-014-9595-3

Maas, E., Gildersleeve-Neumann, C., Jakielski, K., Kovacs, N., Stoeckel, R., Vradelis, H. & Welsh, M. (2019). Bang for your buck: A single-case experimental design study of practice amount and distribution in treatment for childhood apraxia of speech. *Journal of Speech, Language, and Hearing Research.* doi:10.1044/2019_JSLHR-S-18-0212

Maas, E., Gildersleeve-Neumann, C. E., Jakielski, K. J., & Stoekel, R. (2014). Motor-based intervention protocols in treatment of childhood apraxia of speech (CAS). *Current Developmental Disorders Reports, 1,* 197-206. doi:10.1007/s40474-014-0016-4

Maas, E., Robin, D. A., Austermann Hula, S. N., Freedman, S. E., Wulf, G., Ballard, K. J., & Schmidt, R. A. (2008). Principles of motor learning in treatment of motor speech disorders. *American Journal of Speech-Language-Pathology, 17,* 277-298.

McCormack, J., McLeod, S., McAllister, L., & Harrison, L. (2010). My speech problem, your listening problem, and my frustration: The experience of living with childhood speech impairment. *Language, Speech, and Hearing Services in Schools, 41*(4), 379-392. doi:10.1044/0161-1461(2009/08-0129)

Meindl, P., Yu, A., Galla, B.M., Quirk, A., Haeck, C., Goyer, J.P., Lejuez, C.W., D'Mello, S.K., & Duckworth, A.L. (2018). A brief behavioral measure of frustration tolerance predicts academic achievement immediately and two years later. *Emotion.* Advance online publication. doi:10.1037/emo0000492

Merriam-Webster. (2021). Frustrated. In *Merriam-Webster.com dictionary.* Retrieved April 5, 2021 from https://www.merriam-webster.com/dictionary/frustrated.

Miller, G. J., Lewis, B., Benchek, P., Freebairn, L., Tag, J., Budge, K., Iyengar, S. K., Voss-Hoynes, H., Taylor, H. G., & Stein, C. (2019). Reading outcomes for individuals with histories of suspected childhood apraxia of speech. *American Journal of Speech-Language Pathology*, *28*(4), 1432-1447. doi:10.1044/2019_AJSLP-18-0132

Murray, E., Iuzzini-Seigel, J., Maas, E., Terband, H., & Ballard, K. J. (2021). Differential diagnosis of childhood apraxia of speech compared to other speech sound disorders: A systematic review. *American Journal of Speech-Language Pathology*, *30* (1), 279-300.

Murray, E., McCabe, P., & Ballard, K. J. (2014). A systematic review of treatment outcomes for children with childhood apraxia of speech. *American Journal of Speech-Language Pathology*, *23* (3), 486-504.

Murray E., McCabe, P., & Ballard, K. J. (2015). A randomized controlled trial for children with childhood apraxia of speech comparing rapid syllable transition treatment and the Nuffield dyspraxia programme – third edition. *Journal of Speech, Language, and Hearing Research, 58*(3), 669-86. doi:10.1044/2015_JSLHR-S-13-0179

Namasivayam, A. K., Pukonen, M., Goshulak, D., Hard, J., Rudzicz, F., Rietveld, T., Maassen, B., Kroll, R., & van Lieshout, P. (2015). Treatment intensity and childhood apraxia of speech. *International Journal of Language & Communication Disorders, 50*(4), 529–546. doi:10.1111/1460-6984.12154

Rao, R., & Androulakis, I. P. (2019). The physiological significance of the circadian dynamics of the HPA axis: interplay between circadian rhythms, allostasis and stress resilience. *Hormones and Behavior, 110*, 77-89.

Rardin, D. R., & Moan, C. E. (1970). Frustration tolerance and college grade point average. *Perceptual and Motor Skills, 31*(3), 1003–1006. doi:10.2466/pms.1970.31.3.1003

Rusiewicz, H. L., Maize, K., & Ptakowski, T. (2016). Parental experiences and perceptions related to childhood apraxia of speech: Focus on functional implications. *International Journal of Speech-Language Pathology*, *20*(5), 569-580. doi:10.1080/17549507.2017.1359333

Schlotz, W., Hellhammer, J., Schulz, P., & Stone, A. A. (2004). Perceived work overload and chronic worrying predict weekend–weekday differences in the cortisol awakening response. *Psychosomatic Medicine*, *66* (2), 207-214.

Shriberg, L. D., Aram, D. M., & Kwiatkowski, J. (1997). Developmental apraxia of speech: I. Descriptive and theoretical perspectives. *Journal of Speech, Language, and Hearing Research, 40,* 273-285.

Solkoff, N., Gibson, T. A., & Screven, C. G. (1964). Effects of frustration on perceptual-motor performance. *Child Development, 35*(2), 569-585. doi:10.2307/1126720

Strand E. A. (2020). Dynamic temporal and tactile cueing: A treatment strategy for childhood apraxia of speech. *American Journal of Speech-Language Pathology, 29*(1), 30–48. doi:10.1044/2019_AJSLP-19-0005

Thoonen, G., Maassen, B., Wit, J., Gabreëls, F., & Schreuder, R. (1996). The integrated use of maximum performance tasks in differential diagnostic evaluations among children with motor speech disorders. *Clinical Linguistics and Phonetics*, *10*(4), 311-336. doi:10.3109/02699209608985178

Wren, Y. E., Roulstone, S. E., & Miller, L. L. (2012). Distinguishing groups of children with persistent speech disorder: Findings from a prospective population study. *Logopedics Phoniatrics Vocology, 37*(1), 1-10. doi:10.3109/14015439.2011.625973