Copyright 2017 Mariana Aparicio Betancourt

ENVIRONMENTAL INFLUENCES ON COMMUNICATION DEVELOPMENT: IMPLICATIONS FOR CHILDREN WITH NEURODEVELOPMENTAL COMMUNICATION IMPAIRMENTS

BY

MARIANA APARICIO BETANCOURT

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Neuroscience
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2017

Urbana, Illinois

Doctoral Committee:

Associate Professor Laura DeThorne, Chair
Associate Professor Stephanie Ceman
Professor Kara Federmeier
Professor Stephen A. Petrill, The Ohio State University

ABSTRACT

At the intersection of clinical neuroscience and communication sciences and disorders, this dissertation provides a compilation of studies aimed at examining contextual influences on children's communication development and the implications of this work for children with neurodevelopmental communication impairments. As discussed in Chapter 1, the present work is grounded in dynamic systems theory of development and a distributed model of communication, which together emphasize development as a context-dependent dynamic multilevel system that unfolds over time and is shaped by a multitude of factors. Neurodevelopmental communication impairments such as speech sound disorder, language disorder, and autism spectrum disorder affect approximately 1.5 – 16% of children, and are associated with academic, socioemotional, and behavioral difficulties. The work in Chapter 2 directly examines a common form of environmental support for children with neurodevelopmental communication impairments, speech-language therapy. More specifically, it assesses the effectiveness of a multimodal, integrated speech-language intervention in facilitating multisyllabic productions in six children 2-4 years of age with various neurodevelopmental disabilities. It uses single-case and within-subject experimental designs to understand individual trajectories and shape clinical practice. As a complement to the behavioral intervention, Chapter 3 of this thesis explores the novel use of noninvasive biosensors to measure electrical conductance across the skin during speech-language and occupational therapy as a potential support for communication in eight children, ages 2-11, with neurodevelopmental disabilities. Skin conductance is mediated by sympathetic cholinergic sudomotor nerve fibers and has been used extensively in the study of psychological states and processes. However, traditionally its use has been limited to highly controlled laboratory settings, whereas the use of such technology within the context of daily activities remains a major challenge. Next, as a means to examine a broader range of environmental influences, Chapter 4 uses a longitudinal monozygotic (MZ) twin difference method, a genetically sensitive design, to examine four candidate nonshared environmental influences on children's language development: birthweight, breastfeeding, and home reading exposure and parenting (*M* age = 7). This study aims to identify nonshared environmental effects on later language development, at mean ages 10 (*n* = 115 pairs) and 12 years (*n* = 108 pairs), across two assessment contexts: standardized testing and narrative language sampling. Finally, Chapter 5 concludes this dissertation by highlighting the need to study a broader range of contextual factors influencing communication development and its associated mechanisms, incorporate diverse and complementary methodologies, and develop effective communication supports for children with neurodevelopmental communication impairments.

ACKNOWLEDGMENTS

I am grateful and humbled for the guidance and support so many bright scholars, friends, and family members provided me throughout my PhD trajectory. I would first like to thank my mentor and adviser, Dr. Laura DeThorne, for the opportunity to work with her, for striving to incorporate diverse methods, ideas, and cultures, and for her contagious dedication and keen intellect. I would also like to thank her for truly supporting my professional interests and for allowing me to express my own intellectual creativity. I am grateful to my committee members for their thoughtful input, for welcoming me into their laboratories and group meetings, and for their exemplary service to the academic community. I am also indebted to my collaborators in the skin conductance work for their enthusiasm and contributions. My other academic mentors, at and outside of the University of Illinois, have also been instrumental to my development as a scientist.

I would also like to acknowledge the graduate and undergraduate assistants in the Child Language and Literacy Lab who helped make this work possible. Similarly, I owe a special word of gratitude to all participating children and their families. I would also like to thank my colleagues and friends – particularly those with whom I shared long work days – for creating a pleasant and stimulating work atmosphere, and for helping me stay sane with words of encouragement, tea/coffee breaks, and outdoor walks. Thanks as well to acquaintances, or even strangers, who offered a smile when I needed it the most.

Finally, and most importantly, I would like to thank my family, especially my grandparents, siblings, and parents for boundless love and support. To my parents, Adriana and Gonzalo, thank you for filling me with patience and perseverance, and for nurturing my passionate spirit and inquisitive mind. *Gracias, Ma, por brindarme compañía y por nuestras conversaciones eternas que, aunque estuvieras a miles de kilómetros, me hacían sentir tu amor y apoyo incondicional. Pa, como es costumbre: Voluntad, fortaleza, esperanza, sabiduría, unión, prosperidad, mucha fe y mucho amor. Muchas gracias, Currucos.* Thanks as well to my beautiful partner and fellow scientist, Enric, for your support, loving presence, and words of wisdom. *Merci, bonic.*

TABLE OF CONTENTS

CHAPTER 1: Introduction 1

CHAPTER 2: Multimodal Speech-Language Intervention in Children with Neurodevelopmental Communication Impairments: A Within-Participant Study 12

CHAPTER 3: Skin Conductance as an *In-Situ* Marker for Emotional Arousal in Children with Neurodevelopmental Communication Impairments: Methodological Considerations and Clinical Implications 46

CHAPTER 4: Nonshared Environmental Influences on Language Development: A Monozygotic Twin Differences Study 89

CHAPTER 5: Conclusion 161

APPENDIX A: Social Story Akin to That Used as a Desensitization Strategy during the SL-EDA Study 168

APPENDIX B: PARCHISY Cooperative-Play Tasks 171

APPENDIX C: MZ Difference Correlations in Language Outcomes and Environmental Measures for the Full Samples and for Subsamples of Boys and Girls 172

CHAPTER 1:

Introduction

1.1 Communication Development

Although highly complex, communication development appears seemingly effortless to many children. Everyday exchanges such as specifying the source of a tooth ache or inquiring as to the whereabouts of a favorite toy are often taken for granted. Communication development is key to social interaction, later school readiness, emotional wellbeing, and overall participation in society (Forget-Dubois et al., 2009; Qi & Kaiser, 2004; Redmond & Rice, 1998).

It is alleged that communication may begin in utero. Previous findings demonstrate fetuses (37-38 weeks gestational age) are able to recognize familiar rhymes and voices, as measured by differential heart rate changes (Decasper, Lecanuet, Busnel, Granier-Deferre, & Maugeais, 1994; Kisilevsky et al., 2003). Anecdotally, fetuses have also been reported to respond to their mothers touch and voice. Fetal-maternal communication is of course delicately and intrinsically intertwined via their physiology. Postnatally, infants soon begin to communicate their affective states and mothers are able to learn to perceive the differences between an *I'm hungry* cry and an *I'm upset* cry (e.g., Zeskind & Marshall, 1988). During the first three months of life, infants will also communicate by making pleasure sounds, smiling, and by their ability to perceive familiar voices as soothing when in distress. By 12 months many babies have uttered their first word or sign, understand simple directions, (e.g., *come here*), use gestures to communicate, engage in social-sensory routines, and manifest an emerging understanding of joint attention (American Speech-Language-Hearing Association [ASHA], n.d.).

Throughout the lifespan, communication is inherently multimodal, incorporating manual (e.g., gestures, sign), spoken, and written forms of communication (DeThorne & Miller, 2014; see Goldin-Meadow, 2014 for a perspective on the manual modality). During in-person social interactions we communicate both verbally and nonverbally. Although we may consciously attend more to the words and/or signs used when communicating, research shows body language and tone of voice are also key to communication (e.g., Mehrabian & Ferris, 1967; Mehrabian & Wiener, 1967). The multimodal nature of communication often becomes most transparent when specific modalities are impaired. For children with limited speaking abilities, augmentative and alternative communication (AAC) as a natural form of support is key to their development as competent

communicators. In fact, use of AAC has been shown to support diversity of communicative functions (e.g. requesting, commenting) (e.g., Lilienfeld & Alant, 2005), speech and language development (Millar, Light, & Schlosser, 2006; Shane et al., 2012), and decrease challenging behaviors (Mirenda, 1997).

1.2 Neurodevelopmental Communication Impairments

The present research focuses primarily on pediatric neurodevelopmental communication impairments, including speech sound disorder, language disorder, and autism spectrum disorder. Such neurodevelopmental communication impairments may be relatively isolated, as in the case of late-talking toddlers or specific language impairment, or be associated with other forms of cognitive disability (e.g., Intellectual disability). Approximately 1 in 6 children have a neurodevelopmental disability (Boyle et al., 2011), and the prevalence of specific neurodevelopmental communication delays or impairments has been reported as high as 25% (Law, Boyle, Harris, Harkness, & Nye, 2000): the median prevalence of speech-sound delay or disorder ranges from 2-15% in children ages 5-7, with prevalence decreasing with age (Law et al., 2000; McKinnon, McLeod, & Reilly, 2007; Shriberg, Tomblin, & McSweeny, 1999); the median prevalence of receptive and/or expressive language delay or disorder ranges from 3-16% in children ages 2-7, with prevalence decreasing with age (Beitchman, Nair, Clegg, & Patel, 1986; Law et al., 2000; Tomblin et al., 1997); and prevalence of autism in 8-year old children is approximately 1.5% (Christensen et al., 2016). Moreover, males are reportedly more likely than females to have a neurodevelopmental communication impairment across a range of specific diagnoses (Christensen et al., 2016; Law et al., 2000; Shriberg et al., 1999; cf. Beitchman et al., 1986). The etiology of neurodevelopmental communication impairments can be associated with specific genetic conditions (e.g., Fragile X syndrome, Rett syndrome) or environmental factors (e.g. prenatal cytomegalovirus infection, fetal alcohol exposure) but is often considered multifactorial (Kraft & DeThorne, 2014; Rogers, Nulty, Aparicio Betancourt, & DeThorne, 2015).

1.3 Contextual Factors Influencing Communication Development

Examining the contextual factors influencing communication development will provide us with a better understanding of the causal influences on communication development and the supports needed for children with neurodevelopmental disabilities. More specifically, the present

work focuses on the overarching question of what contextual factors shape children's communication development and understanding the process by which they do so. This inquiry has been grounded in the dynamic systems theory of development (Lerner, 2006; Oyama, Griffiths, & Gray, 2001; Smith & Thelen, 2003; Thelen, 2005; Thelen & Smith, 2006) in partnership with a distributed model of communication (DeThorne & Miller, 2014; Hengst, 2015), which together have led me to conceptualize communication as a) embedded within a dynamic multilevel system, b) unfolding over time, and c) as context dependent.

1.3.1 Communication development within a dynamic multilevel system. Communication development occurs within a dynamic multilevel system (e.g., from molecular to cultural), influenced by multiple causal factors. For example, initial synaptic contacts are thought to be genetically determined, but the environment in which the child develops will play a major role in stabilizing these contacts in order to form the neural networks that support specific functions such as communication. Early neuronal networks are undoubtedly shaped by individual genetic differences (e.g., Fragile X syndrome), as well as environmental factors such as explicit trauma (e.g., periventricular white matter lesion) and the nature of early caregiver communication practices. At a cultural-historical level, histories of immigration and discrimination have shaped everything from epigenetic influences on health to community attitudes regarding language and dialect differences. Multiple factors across different levels will continuously interact with one another to influence communication development.

Chapter 2 of the present dissertation specifically explores one form of common environmental support for children with neurodevelopmental communication impairments, speech-language therapy. Specifically, this study capitalized on the multimodal nature of communication on the development of children's multisyllabic speech productions using a within-participant design. The variety of strategies used, including play-based strategies, and the incorporation of VocSyl, a novel computerized software program, targeted multiple facets of the children's communication system. Through the use of within-participant methods, this study controlled for many potential influences on child language use (e.g., genetic predisposition, pre-and perinatal factors, parenting) while measuring the effect of intervention aside from maturation.

Whereas Chapter 2 relied exclusively on behavioral measures associated with intervention outcomes, Chapter 3 highlights data from two proof of concept intervention studies that specifically integrated a physiological measure of arousal, skin conductance, as a relevant outcome

measure related to behavioral intervention (Aparicio Betancourt, DeThorne, Karahalios, & Kim, 2017). This interdisciplinary work across neuroscience, speech and hearing science, and computer science represents the first published manuscript to share data on *in situ* skin conductance assessment during intervention in children with communication impairments and consequently focuses on guidelines for future research. Given that many children with neurodevelopmental communication disabilities process sensory information differently and may have difficulty expressing their experience, there is long-term potential for skin conductance to be used as a form of environmental support to facilitate communication between child and caregiver.

Finally, Chapter 4 builds explicitly on my prior work on environmental effects (Harlaar, DeThorne, Mahurin-Smith, Aparicio Betancourt, & Petrill, 2016; Rogers et al., 2015) by examining potential nonshared environmental influences on language development, specifically peri-, early post-natal, and socio-linguistic factors, while controlling for genetics through use of a monozygotic twin difference design.

1.3.2 Communication development unfolds over time. Communication can also be perceived as a system with nested processes that unfold over many timescales (e.g., from milliseconds to centuries). Communication development is thus composed of many individual elements that are continuously exposed to complex environmental influences; no single element has causal priority in explaining behavior or its changes (Thelen & Smith, 2006). Development is shaped by the dynamic and complex interplay between various causal influences over time. Whereas neural excitation occurs in a matter of milliseconds, emergence of many speech-language milestones take months or years to emerge. In addition, distributed theories of communication stress that patterns of interaction emerge over time as individuals interact with each other and their specific contexts (DeThorne & Miller, 2014; Hengst, 2015).

The complex interplay between environmental and genetic factors over time has been conceptualized through gene-environment correlations, gene-environment interactions (GxE), and epigenetics. Gene-environment correlations describe how genetic and environmental influences are often interdependent (Scarr & McCartney, 1983). For example, a child genetically predisposed to be loquacious is likely to elicit more verbal interactions, which in turn may offer more practice and exposure to language-rich interactions, an effect referred to as evocative gene-environment correlation. In addition, genetic differences can moderate environmental effects. For example, breastfeeding appears to confer benefits in IQ for some children but not others, depending in part

on allelic variation (Caspi et al., 2007). Finally, epigenetics refers to processes that induce changes in gene expression without altering the genotype. Epigenetic regulation of genes, via e.g. histone modification, has been shown to lead, in some cases, to profound changes in phenotype. For example, although the epigenomes of monozygotic twins are originally indistinguishable, their environments become increasingly differentiated as they age and so do their epigenomes (Fraga et al., 2005). The interaction of multiple influences at any point in time may alter gene expression and provides a possible explanation for high discordance rates between monozygotic twins. Environmental influences such as diet and parental interactions have been associated with modulation of gene expression (Verduci et al., 2014; Weaver et al., 2004). To illustrate, epigenomic alterations at a glucocorticoid receptor gene promoter in the hippocampus as a result of maternal nurturing of rat pups have been reported, in which high vs. low nurturing of the pups during their first week after birth resulted in calm vs. anxious rats (Weaver et al., 2004). Depending on context, either behavioral phenotype may be deemed advantageous. Over time, all of our experiences will shape us and our interaction with the environment; they may also shape our children and our children's children.

Although the interplay between genetic and environmental influences on language is in its infancy (Dale, Tosto, Hayiou-Thomas, & Plomin, 2015; Rogers et al., 2015), studying communication development over time is a hallmark of behavioral research, though rarely within a genetically-sensitive design. The single-case multiple probe across behaviors experimental design presented in Chapter 2 offers the strength of comparing behaviors (in this case treated v. untreated speech targets) within the same children, therefore eliminating genetic variance as a confound. The data reported within Chapter 3 is also based on two studies utilizing single-case design, and time-series analyses. Both chapters focus on intense data collection procedures, collected across 15 to 41 weeks of intervention, with careful experimental control for influences like maturation and spurious environmental effects. Finally, chapter 4 explicitly capitalizes on a rich longitudinal data set of 216 - 230 identical twins from the Western Reserve Reading and Math Project (Petrill, Deater-Deckard, Thompson, DeThorne, & Schatschneider, 2006) that provides a unique genetically sensitive design and considers language outcomes at two separate time points two years apart.

1.3.3 Communication is context dependent. The final principle emphasized in the present work is that communication is context dependent, functioning within a myriad of environmental

constraints and supports. Specifically, communication is conceptualized as distributed across modalities, partners, and embedded within specific activities, situated within specific settings, and within a specific cultural-historical context (DeThorne & Miller, 2014; Hengst, 2015). Ultimately, one's ability to engage in successful communication exchanges is context-dependent. For example, how one interacts with the environment and communicates with others will depend on who you are communicating with, and the supports in place that allow you to communicate in the first place. Caregivers of many of the children we worked with would often translate their children's use of words and/or gestures embedded within specific activities, highlighting how communication is distributed across partners and shaped over time. Caregivers would describe how the child was referring to a specific experience or motivating us to engage in a social interaction they particularly enjoyed. Without support from a familiar communication partner, we would not have been able to understand the full extent of the child's message. For example, when first introducing the biosensors, which are watch-like bands with embedded electrodes, to measure skin conductance in one of the children during speech-language therapy (Ch 3), the child immediately responded by saying *papa wa*. The child's mom explained that he was referring to a particular social routine associated with his father's watch. Familiarity between communication partners breeds understanding and successful communication is dependent on context and shared history built over time.

A common theme across the studies presented here is understanding what contextual factors shape and support communication development. The intervention study presented in Chapter 2 focuses on the development of communication practices through shared activities and multimodal supports. Specifically, the same therapists developed routines with each child around child-centered play with a familiar set of objects. In addition, although the dependent variable was speech production, the intervention supported overall communication through use of all modalities and alternative and augmentative communication supports, including specific signs (e.g., *open*, *more*) and access to a mid-tech speech-generating device (i.e., GoTalk 20+). Chapter 3 explicitly explores novel use of technology (i.e., biosensors) as a contextual support for communication across children with neurodevelopmental communication impairments and their communication partners. Specifically, this chapter offers guidelines as well as potential challenges associated with the use of this technology within such contexts. Finally, Chapter 4, explicitly uses a genetically-sensitive design (i.e., MZ-difference method) to assess the extent to which specific perinatal, early

postnatal, social, and linguistic contexts are associated with differences in children's language outcomes. In addition, it examines language outcomes within two different assessment contexts (narrative language sampling and standardized testing), acknowledging the potential influence of context on a child's performance reported in prior work (DeThorne et al., 2008; Harlaar et al., 2016).

***

In summary, communication develops within a dynamic multilevel system over time. As summarized by Thelen (2005), "every act in every moment is the emergent product of context and history, and no component has causal priority" (p. 271). It is influenced by the continuous interaction of various contextual factors including communication partners, modalities, specific activities, and more generally by the interplay between environmental and genetic influences. Altogether, the present work highlights the need for the development of effective communication supports for children with neurodevelopmental communication impairments, and the need to widen the lens through which we view contextual factors influencing communication development.

References

American Speech-Language-Hearing Association [ASHA] (n.d.). *How does your child hear and talk?* Retrieved March 18, 2017 from http://www.asha.org/public/speech/development/chart/.

Aparicio Betancourt, M., DeThorne, L. S., Karahalios, K. & Kim, J. G. (2017). Skin conductance as an *in situ* marker for emotional arousal in children with neurodevelopmental communication impairments: Methodological considerations and clinical implications. *ACM Transactions on Accessible Computing*, *9*(3), 8. http://dx.doi.org/10.1145/3035536

Beitchman, J. H., Nair, R., Clegg, M., & Patel, P. G. (1986). Prevalence of speech and language disorders in 5-year-old kindergarten children in the Ottawa-Carleton region. *Journal of Speech and Hearing Disorders*, *51*(2), 98-110. doi:10.1044/jshd.5102.98

Boyle, C. A., Boulet, S., Schieve, L. A., Cohen, R. A., Blumberg, S. J., Yeargin-Allsopp, M., ... & Kogan, M. D. (2011). Trends in the prevalence of developmental disabilities in US children, 1997–2008. *Pediatrics*, *127*(6), 1034-1042. doi: 10.1542/peds.2010-2989

Caspi, A., Williams, B., Kim-Cohen, J., Craig, I. W., Milne, B. J., Poulton, R., ... & Moffitt, T. E. (2007). Moderation of breastfeeding effects on the IQ by genetic variation in fatty acid metabolism. *Proceedings of the National Academy of Sciences*, *104*(47), 18860-18865. doi:10.1073/pnas.0704292104

Christensen, D. L., Baio, J., Braun, K. V., Bilder, D., Charles, J., Constantino, J. N., ... & Yeargin-Allsopp, M. (2016). Prevalence and characteristics of autism spectrum disorder among children aged 8 years — Autism and Developmental Disabilities Monitoring Network, 11 sites, United States, 2012. *MMWR Surveillance Summaries*, *65*(3), 1-23. http://dx.doi.org/10.15585/mmwr.ss6503a1

Dale, P. S., Tosto, M. G., Hayiou-Thomas, M. E., & Plomin, R. (2015). Why does parental language input style predict child language development? A twin study of gene–environment correlation. *Journal of communication disorders*, *57*, 106-117. http://dx.doi.org/10.1016/j.jcomdis.2015.07.004

DeCasper, A. J., Lecanuet, J. P., Busnel, M. C., Granier-Deferre, C., & Maugeais, R. (1994). Fetal reactions to recurrent maternal speech. *Infant behavior and development*, *17*(2), 159-164. http://dx.doi.org/10.1016/0163-6383(94)90051-5

DeThorne, L. S., Petrill, S. A., Hart, S. A., Channell, R. W., Campbell, R. J., Deater-Deckard, K., ... & Vandenbergh, D. J. (2008). Genetic effects on children's conversational language use. *Journal of Speech, Language, and Hearing Research*, *51*(2), 423-435. doi:10.1044/1092-4388(2008/031)

DeThorne, L.S., & Miller, C. A. (2014). Communication development: Distributed across people, resources and time. In J. Nussbaum (Ed.), *Handbook of lifespan communication* (53-70). New York, New York, USA: Peter Lang Publishing

Forget-Dubois, N., Dionne, G., Lemelin, J. P., Pérusse, D., Tremblay, R. E., & Boivin, M. (2009). Early child language mediates the relation between home environment and school readiness. *Child development*, *80*(3), 736-749. doi:10.1111/j.1467-8624.2009.01294.x

Fraga, M. F., Ballestar, E., Paz, M. F., Ropero, S., Setien, F., Ballestar, M. L., ... & Boix-Chornet, M. (2005). Epigenetic differences arise during the lifetime of monozygotic twins. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(30), 10604-10609. doi: 10.1073/pnas.0500398102

Goldin-Meadow, S. (2014). Widening the lens: what the manual modality reveals about language, learning and cognition. *Phil. Trans. R. Soc. B, 369*(1651), 20130295. http://dx.doi.org/10.1098/rstb.2013.0295

Harlaar, N., DeThorne, L. S., Mahurin-Smith, J., Aparicio Betancourt, M., & Petrill, S. A. (2016). Longitudinal Effects on Early Adolescent Language: A Twin Study. *Journal of Speech, Language, and Hearing Research*, *59*(5), 1059-1073. doi:10.1044/2016_JSLHR-L-15-0257

Hengst, J. A. (2015). Distributed communication: Implications of cultural-historical activity theory (CHAT) for communication disorders. *Journal of communication disorders*, *57*, 16-28. http://dx.doi.org/10.1016/j.jcomdis.2015.09.001

Kisilevsky, B. S., Hains, S. M., Lee, K., Xie, X., Huang, H., Ye, H. H., ... & Wang, Z. (2003). Effects of experience on fetal voice recognition. *Psychological science*, *14*(3), 220-224. https://doi.org/10.1111/1467-9280.02435

Kraft, S. J., & DeThorne, L. S. (2014). The brave new world of epigenetics: Embracing complexity in the study of speech and language disorders. *Current Developmental Disorders Reports*, *1*(3), 207-214. doi:10.1007/s40474-014-0024-4

Law, J., Boyle, J., Harris, F., Harkness, A., & Nye, C. (2000). Prevalence and natural history of primary speech and language delay: Findings from a systematic review of the literature. *International Journal of Language and Communication Disorders*, *35*(2), 165-188.

Lerner, R. M. (2006). Developmental science, developmental systems, and contemporary theories of human development. In W. Damon & R.M. Lerner (Eds.), *Handbook of Child Psychology* (Vol.1), (1-17). Hoboken, NJ, USA: John Wiley & Sons. doi:10.1002/9780470147658.chpsy0101

Lilienfeld, M., & Alant, E. (2005). The social interaction of an adolescent who uses AAC: The evaluation of a peer-training program. *Augmentative and Alternative Communication*, *21*(4), 278-294. http://dx.doi.org/10.1080/07434610500103467

McKinnon, D. H., McLeod, S., & Reilly, S. (2007). The prevalence of stuttering, voice, and speech-sound disorders in primary school students in Australia. *Language, Speech, and Hearing Services in Schools*, *38*(1), 5-15. doi:10.1044/0161-1461(2007/002)

Mehrabian, A., & Ferris, S. R. (1967). Inference of attitudes from nonverbal communication in two channels. *Journal of consulting psychology*, *31*(3), 248. http://psycnet.apa.org/doi/10.1037/h0024648

Mehrabian, A., & Wiener, M. (1967). Decoding of inconsistent communications. *Journal of Personality and Social Psychology,* 6(1), 109-114. http://psycnet.apa.org/doi/10.1037/h0024532

Millar, D. C., Light, J. C., & Schlosser, R. W. (2006). The impact of augmentative and alternative communication interventions on the speech production of individuals with developmental disabilities: A research review. Journal of Speech, Language, and Hearing Research, 49, 248-264. doi:10.1044/1092-4388(2006/021)

Mirenda, P. (1997). Supporting individuals with challenging behavior through functional communication training and AAC: Research and review. Augmentative and Alternative Communication, 13, 207-225. Retrieved from http://informahealthcare.com/loi/aac/

Oyama, S., Griffiths, P. E., & Gray, R. D. (2001). Introduction: What is developmental systems theory. In S. Oyama, P. E. Griffiths, & R. D. Gray (Eds.), *Cycles of contingency: Developmental systems and evolution* (pp. 1-11). Cambridge, Massachusetts, USA: MIT Press.

Petrill, S. A., Deater-Deckard, K., Thompson, L. A., DeThorne, L. S., & Schatschneider, C. (2006). Reading skills in early readers genetic and shared environmental influences. *Journal of Learning Disabilities*, *39*(1), 48-55. doi:10.1177/00222194060390010501

Qi, C. H., & Kaiser, A. P. (2004). Problem behaviors of low-income children with language delays: An observation study. *Journal of Speech, Language, and Hearing Research,* 47(3), 595-609. doi: 10.1044/1092-4388(2004/046).

Redmond, S. M., & Rice, M. L. (1998). The socioemotional behaviors of children with SLI: Social adaptation or social deviance? *Journal of Speech, Language, and Hearing Research,* 41(3), 688-700. doi: 10.1044/jslhr.4103.688.

Rogers, C. R., Nulty, K. L., Aparicio Betancourt, M., & DeThorne, L. S. (2015). Causal effects on child language development: A review of studies in communication sciences and disorders. *Journal of communication disorders*,*57*, 3-15. http://dx.doi.org/10.1016/j.jcomdis.2015.06.004

Scarr, S., & McCartney, K. (1983). How people make their own environments: A theory of genotype→ environment effects. *Child development*, 424-435. http://www.jstor.org/stable/1129703

Shane, H. C., Laubscher, E. H., Schlosser, R. W., Flynn, S., Sorce, J. F., & Abramson, J. (2012). Applying technology to visually support language and communication in individuals with autism spectrum disorders. *Journal of autism and developmental disorders*, *42*(6), 1228-1235. doi:10.1007/s10803-011-1304-z

Shriberg, L. D., Tomblin, J. B., & McSweeny, J. L. (1999). Prevalence of speech delay in 6-year-old children and comorbidity with language impairment. *Journal of speech, language, and hearing research*, *42*(6), 1461-1481. doi:10.1044/jslhr.4206.1461

Smith, L. B., & Thelen, E. (2003). Development as a dynamic system. *Trends in cognitive sciences*, *7*(8), 343-348. http://dx.doi.org/10.1016/S1364-6613(03)00156-6

Thelen, E. (2005). Dynamic systems theory and the complexity of change. *Psychoanalytic Dialogues*, *15*(2), 255-283. http://dx.doi.org/10.1080/10481881509348831

Thelen, E., & Smith, L. B. (2006). Dynamic systems theories. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology* (Vol. 1), (258-312). Hoboken, NJ, USA: John Wiley & Sons.

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research*, *40*(6), 1245-1260. doi:10.1044/jslhr.4006.1245

Verduci, E., Banderali, G., Barberi, S., Radaelli, G., Lops, A., Betti, F., ... & Giovannini, M. (2014). Epigenetic effects of human breast milk. *Nutrients*, *6*(4), 1711-1724. doi:10.3390/nu6041711

Weaver, I. C., Cervoni, N., Champagne, F. A., D'Alessio, A. C., Sharma, S., Seckl, J. R., ... & Meaney, M. J. (2004). Epigenetic programming by maternal behavior. *Nature neuroscience*, *7*(8), 847-854. doi:10.1038/nn1276

Zeskind, P. S., & Marshall, T. R. (1988). The relation between variations in pitch and maternal perceptions of infant crying. *Child Development*, 193-196. doi:10.2307/1130401

CHAPTER 2:

Multimodal Speech-Language Intervention in Children with Neurodevelopmental Communication Impairments: A Within-Participant Study

Abstract

The present study contributes to the science of the individual by utilizing within-participant data to examine the effectiveness of a multimodal, integrated speech-language intervention. It offers a direct follow-up to experimental group data presented in DeThorne, Aparicio Betancourt, Karahalios, Halle, and Bogue (2015), and provides a unique opportunity to examine the benefits and limitations of analyzing individual outcomes relative to group outcomes within the same data set. Specifically, this study focused on six children (2-4 years) with neurodevelopmental communication impairments at the single-word stage of development. The multimodal intervention aimed to facilitate multisyllabic speech targets using a combination of motor practice, via use of the computerized feedback software VocSyl, and developmental play. Thirty multisyllabic speech targets were selected for each child, 15 of which were treated, and 15 served as control targets. Consistent with single-case multiple probe across behaviors experimental design, the 15 treatment targets were introduced five at a time, as a means to compare change in treated relative to untreated targets. Converging forms of within-participant evidence provided support for a positive treatment effect. Specifically, all six children performed better on treatment targets relative to control targets at post-treatment and maintenance, and data from parent report indicated gains in expressive multisyllabic vocabulary inventories following intervention. However, single-case experimental data was able to establish a clear functional relationship between the intervention and increased multisyllabic targets production in only half of the participants (3/6). This study provides converging support for a multimodal, integrated speech-language intervention focused on facilitating multisyllabic productions, and highlights the particular benefits and challenges of focusing on individual outcomes for the study of children's speech-language development.

2.1 Introduction

The purpose of the present study was to examine the effectiveness of a multimodal speech-language intervention using within-participant methodology. The intervention focused explicitly on children's ability to combine syllables, either as word combinations (e.g., rocking chair) or as a multisyllabic word (e.g., butterfly). The ability to produce multisyllabic words and phrases represents an important developmental milestone, typically reached around 18-24 months of age, that integrates the motor processes of speech sound production with the semantic and grammatical aspects of language. Many children with neurodevelopmental communication impairments, including autism, demonstrate difficulties in both speech and language domains, including delays in the development of multisyllabic productions (Highman, Hennessey, Sherwood, & Leitão, 2008; Law, Boyle, Harris, Harkness, & Nye, 2000; Preston & Edwards, 2007; Tager-Flusberg et al., 2009). Despite the prevalence of children with concomitant speech and language needs, interventions explicitly aimed at facilitating multisyllabic production are scarce and often focus on one domain without explicitly addressing the other.

Language-based approaches to facilitating multisyllabic productions focus on using developmental play routines such as social-interaction games and communication temptations to model multisyllabic words and phrases (Woods & Wetherby, 2003; see also Harjusola-Webb & Robbins, 2012). Increased salience of the targets is accomplished through providing a high frequency of verbal models (e.g.,Yosick, Muskat, Bowen, Delfs, & Shillingsburg, 2016), often produced with exaggerated prosody and supplemented with visual referents such as objects or gestures (Girolametto, Pearce, & Weitzman, 1995, 1996; Robertson & Weismer, 1999). When models are provided through augmentative and alternative communication systems (AAC), the practice is commonly referred to as aided language stimulation (Binger & Light, 2007; Iacono, Mirenda, & Beukelman, 1993; Romski & Sevcik, 2003). Models, whether spoken or aided, are sometimes paired with attempts to elicit production through strategies such as forced choice (e.g., "Do you want the *elephant* or the *tiger*?") and carrier phrases (e.g., "You want the …?"). Language-based approaches, though useful in facilitating word use, often do not provide explicit support for motor speech impairments.

In contrast to developmental language-based approaches, speech-based interventions have centered largely on direct imitation of multisyllabic productions in drill-like routines, often modeling targets with slower rate and exaggerated intonation (cf. Wan et al., 2011). Augmented

feedback is also often provided via tactile and visual cues of key articulation movements or phonological features (Strand & Debertine, 2000; Strand, Stoeckel, & Baas, 2006; Velleman, 2002). An articulation cue is one that provides information related to the manner, place, or voicing of a particular speech sound. For instance, if a child produced the word *vacuum* as *vavuum* a clinician might tap a finger on the child's throat to highlight that the middle /k/ sound is made by raising the back of the tongue to make contact with the soft palate. Phonological cues used to facilitate multisyllabic productions provide multimodal information regarding syllable number. For example, a clinician might clap along with each syllable as a word or phrase is being produced. Both traditional static tools (e.g., pacing board; DeThorne, Aparicio Betancourt, Karahalios, Halle, & Bogue, 2015; Kumin, Councill, & Goodman, 1995; Velleman, 1994) and more recently dynamic online feedback tools (e.g., VocSyl; DeThorne et al., 2015) have also been used for making the construct of the syllable more salient. Despite the utility of speech-based approaches in supporting underlying motor speech skills, such approaches often do not support successful communication more broadly. In addition, the potential benefit of AAC use on the motoric development of speech production remains relatively unexamined (cf., DeThorne, Johnson, Walder, & Mahurin-Smith, 2009; Romski et al., 2010, Schlosser & Wendt, 2008).

Although the ability to combine words and syllables represents an important developmental milestone that is delayed or impaired in a variety of clinically-identified populations, evidence to support integrated speech-language treatment practices in this area is relatively sparse. To address this gap, DeThorne and colleagues (2015) utilized a mixed method design to evaluate the effectiveness of two integrated speech-language interventions aimed to facilitate multisyllabic productions in children with speech-language impairments at the single-word stage of development (age 2-8 years). Both interventions integrated language and speech-based strategies, but the traditional condition used a pacing board to emphasize syllable breaks, whereas the other intervention used VocSyl, a computerized feedback tool, to provide a dynamic display of syllable boundaries and other prosodic features. The experimental group design focused on a total of 18 children systematically assigned to either one of the two speech-language intervention conditions (i.e., Pacing Board, VocSyl) or a control condition that emphasized social interaction within a play group setting. Evaluation focused on children's percent accuracy of treated and control multisyllabic targets at two time points: post-intervention and maintenance sessions. In sum, both the speech-language intervention groups produced more multisyllabic treated targets (average gain

of 7 to 9 words) relative to the control group (average gain of 3 to 4 words); however, the only statistically significant between-group difference was between VocSyl and the Control group at the maintenance session. Statistically significant within-group differences between treatment and control targets were also observed for both the speech-language intervention groups, whereas they were not observed for the Control group. The 2015 publication focused primarily on experimental group data which offered an opportunity to assess statistically significant mean differences across and within groups. The overarching treatment study, however, was explicitly designed to assess both group and individual outcomes by merging experimental between-subject, within-subject[1], and single-case[2] methodologies. Consequently, it offers a unique opportunity to highlight the utility of emphasizing individual outcomes relative to group outcomes within the study of children's speech-language development.

Single-case experimental designs, one type of within-participant methodology[3], provide complementary contributions to the knowledge provided by group experimental designs by employing within- and between-subjects comparisons. The development of current single-case research emerged from the behavioral work of Skinner (1904-1990) and gained recognition around the 1950s and 1960s (Kazdin, 2011). The approach has proliferated within the field of special education, where it has been primarily used to examine interventions aimed at either reducing challenging behaviors or increasing desirable ones. The defining element of this experimental methodology is the use of individual cases, often individuals, that serve as their own control to conduct systematic within-case evaluations within and across different conditions/phases (Horner et al., 2005; Kazdin, 2011). For example, a child's observed behavior across multiple data points prior to intervention (i.e., baseline) is compared to that same child's behavior during and/or after intervention. Repeated assessment, within and across conditions (e.g., interventions, behaviors, settings, participants) is critical to documenting an experimental effect. Analyses are traditionally contingent on systematic visual inspection of one or more observable and operationally defined

[1] A within-subject experimental design also known as a repeated-measures design traditionally focuses on group outcomes even though it is also possible to highlight individual outcomes by making comparisons across conditions within the same individual.

[2] Single-case designs also known as single-subject designs refer to a specific experimental methodology, as opposed to correlational or descriptive, that systematically evaluates an intervention on a particular case (e.g., a person, school) and each case serves as its own control by repeatedly measuring the dependent variable(s) over time (e.g. see Horner et al., 2005).

[3] We use the term within-participant methodology to refer to any comparisons of data being made within the same individual that focuses on individual outcomes (i.e., it encompasses both within-subject and single-case designs).

dependent variable(s) graphed over time. An experimental effect is demonstrated when manipulation of the independent variable precedes change in the dependent variable(s). At least three within- or between- case replications of the anticipated intervention effect at different time points are required to demonstrate experimental control and confirm the existence of a causal, or functional relationship between the independent and dependent variables (Horner et al., 2005; Kazdin, 2011).

In addition to their ability to examine causal relationships, single-case designs require fewer participants than experimental groups designs, are more cost-effective, offer the flexibility to tailor interventions to the needs of individual children, and provide the opportunity to provide richer descriptions of individual children. Such advantages are particularly relevant given critiques that the group outcomes used in experimental group designs (e.g., averages) tend to mask individual variability and may have limited value in understanding the needs of individuals. As provocatively stated by Todd Rose (2016), "the moment you need to make a decision about any individual – the average, [often used in group research], is useless. Worse than useless, in fact, because it creates the illusion of knowledge, when in fact the average disguises what is most important about an individual" (p. 11). Given that the emphasis of medicine and education is on the individual patient or student, research that focuses on individual characteristics and outcomes is highly valuable.

Although single-case experimental methodology is capable of helping establish evidence-based interventions (Horner et al., 2005), it is not explicitly discussed within the American Speech-Language-Hearing Association's levels of evidence, with meta-analysis of more than one randomized controlled trial (RCT), a type of between-group design, explicitly advocated as the gold standard (Level Ia) for intervention research (ASHA, n.d.). Although some literature within the field of communication sciences and disorders acknowledges single-case methodology as experimental, it tends to be placed as Level II research below that of RCTs (Justice & Fey, 2004; see also Johnson, 2006). Critiques of single-case methodology have focused primarily on concern about what generalizations can be made from such few cases, and the limitations of this method for studying developmental behaviors like language that do not necessarily revert to baseline conditions once the treatment is withdrawn. External validity of the results is improved through replication of the effects across different studies, participants, settings, or behaviors. Although not all single-case designs are recommended for the study of speech and language, some are useful in studying these developmental processes.

A mixed methodology that combines single-case, within-subject, and between-subject designs offers a unique opportunity to examine the complementary strengths and limitations of focusing on individual versus group outcomes. Accordingly, the present study explicitly examines within-participant data, including single-case data from a mixed method study reported in DeThorne and colleagues (2015). Whereas the 2015 paper focused on experimental group data, the present study offers a unique opportunity to highlight individual outcomes using within-participant data from the same project. Specifically, this particular study focuses on the learning trajectories of the six children who participated in the VocSyl intervention. The specific research questions are as follows:

1. Is there a functional relationship between the implementation of the speech-language intervention and the production of the multisyllabic speech targets?
2. What information does the within-participant data provide relative to the experimental group results published in DeThorne and colleagues (2015)?

2.2 Methods

2.2.1 Participants. Approval for this project was attained through the University of Illinois Institutional Review Board prior to recruitment. This study focused on six children, ages 2 - 4 years, who completed the computerized feedback intervention (VocSyl) in DeThorne et al. (2015). We focus here on children enrolled in the computerized feedback intervention because it revealed the most promising findings based on experimental group data. The larger project, including the six participants in the present study, focused on children at the single-word stage of development with suspected spoken language and oral-motor coordination difficulties who were recruited through key community agencies and professional organizations (e.g., The Autism Program, local school districts, pediatricians, private speech-language therapists, and the University's Speech-Language Clinic). Interested families first participated in a phone screening, followed by an initial assessment at the University's Speech-Language Clinic. During the initial assessment, a speech-language sample was obtained, and children and/or their caregivers were asked to complete or were administered several instruments including: the MacArthur-Bates Communicative Development Inventory: Words and Gestures (CDI; Fenson et al., 2007), the Verbal Motor Production Assessment for Children (VMPAC) Oromotor Production in Word Sequences and Sentences section (Hayden & Square, 1999), and the Ages and Stages Questionnaire of Problem

Solving Skills (Squires & Bricker, 2009). Inclusionary criteria were specified as a) between the ages of 2-8 years; b) at least 30 expressive vocabulary words on a parent report measure (i.e., the CDI: Words and Gestures); c) fewer than 20 communicative, spontaneous multisyllabic productions based on a 20-minute observational speech-language sample; and d) equal to or less than 30% accuracy on the VMPAC Oromotor Production in Word Sequences and Sentences section. All participants' primary language was English. Table 2.1 provides a summary of descriptive data for all six participants included in the present study by pseudonym.

Table 2.1. *Participant Demographics*

	Pyrros	Karis	Santiago	Angelo	Heidi	Conley
Age	2;7	4;10 [a]	3;9	3;0	3;8	2;2
Languages Exposed To	English	English, Russian, Romanian	English, Spanish	English	English	English
Gender	Boy	Boy	Boy	Boy	Girl	Boy
Race/Ethnicity	White	White	White, Hispanic	White	White	White
Relevant Clinical Diagnosis	Unspecified Developmental Delay	Autism	Unspecified Developmental Delay	Speech Sound Disorder	Autism	Unspecified Developmental Delay
Parent Education	19	18	19	17	17	19
CDI-receptive (pre)	130	327	242	349	117	256
CDI-expressive (pre)	82	276	183	154	88	64
ASQ	Monitoring Zone	Typical	Typical	Delay	Delay	Typical
VMPAC	0%	0%	30%	0%	26%	8%
# Spontaneous multisyllabic productions	6	7	6	0	7	3

Parent education refers to the primary caregiver's self-reported number of years of formal education; CDI refers to the MacArthur-Bates communicative development words and gestures receptive and expressive vocabulary inventories pre-intervention; ASQ refers to the Ages and Stages Questionnaire of Problem Solving Skills, and was used as a measure of cognitive development; VMPAC refers to the number of accurately imitated syllables from the Verbal Motor Production Assessment for Children, Oromotor Production in Word Sequences and Sentences section; # Spontaneous multisyllabic productions is based on a 20-minute observational speech-language sample at the initial assessment. [a]Karis' age was reported as 4;8 in Aparicio Betancourt, DeThorne, Karahalios, & Kim, 2017 but should read as 4;10 (i.e., 4.8 years).

2.2.2 Design. To follow-up on experimental group data presented in DeThorne and colleagues (2015), the present study focuses on individual outcomes, specifically including within-subject and a single-case multiple-probe across behaviors design with six participants enrolled in a multimodal speech-language intervention aimed at facilitating multisyllabic productions. Key to the design was the selection of thirty multisyllabic speech targets from the MacArthur-Bates CDI for each child. Preference was given to multisyllabic targets, with 2 or 3 syllables each, within the children's receptive vocabulary, but outside their expressive vocabulary (e.g., *tiger*). In addition, when possible we selected targets that consisted of sounds from within the child's phonetic repertoire, or alternatively, as phonetically simple as possible based on the Index of Phonetic Complexity (IPC; Jakielski, Maytasse, & Doyle, 2006; see also Morris, 2009). When additional targets were needed, mono- or disyllabic words were combined to make multisyllabic targets (e.g., *drink it*, *teddy bear*)[4]. The 30 multisyllabic targets for each child were assigned to one of two 15-target lists; the two lists were randomly assigned to either treatment or control. The targets were balanced between the two lists based on a) syllable number (i.e., two vs. three), b) CDI semantic category (e.g., animals, verbs, adjectives, house items), and c) and phonetic complexity (IPC). Additionally, each list of 15 targets was further divided into three sets of 5, again balanced by number of syllables, semantic category, and IPC rating as best as possible. The three target sets were randomly designated as first, second, and third, and were treated one set at a time for each child.

2.2.3 Procedures. Treatment and assessment sessions were conducted by two graduate students and overseen by an ASHA-certified speech-language pathologist at the Speech-Language Clinic of the University of Illinois. All sessions were video-recorded, with children's attempts at the targets being phonetically transcribed based on the International Phonetic Alphabet. Correct production for an individual target was defined as verbal marking of the appropriate number of syllables, paired with a minimum of 50% of the target phonemes in the correct order.

2.2.3.1 Treatment. Treatment sessions lasted approximately 45-minutes each, and were generally conducted twice/week. The total number of treatment sessions was determined by each child's rate of progress, with a maximum of 12 treatment sessions per five-item target set (for a maximum of 36 treatment sessions). Treatment sessions provided multimodal support and

[4] Only 4% (7/180) of all the speech targets combined across children consisted of word combinations.

combined a) drill-based motor practice and b) child-centered developmental play. During motor practice, the clinicians attempted to elicit from the child each of the five multisyllabic targets within a set five times by providing verbal models using VocSyl for visual feedback. VocSyl is a novel software program designed specifically for this study to provide real-time visual feedback of syllables, rate, pitch, and volume (see Figure 2.1; Hailpern, Harris, La Botz, Birman, & Karahalios, 2012; Hailpern, Karahalios, DeThorne, & Halle, 2010). The clinician modeled each multisyllabic target using exaggerated intonation and a reduced rate while simultaneously tapping the syllable representations on the screen. In addition, supplemental articulation (i.e., visual and tactile) and phonology cues were provided to shape the production as needed (cf. Strand & Debertine, 2000; Strand et al., 2006; Velleman, 2002).

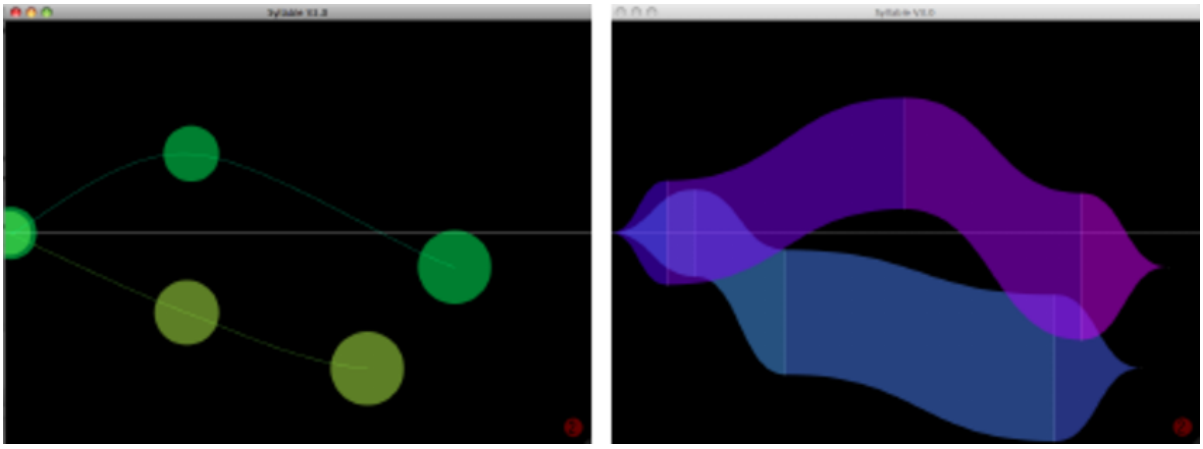


Figure 2.1. Visualization from VocSyl of a three-syllable word (in Hailpern et al., 2010 reprinted with permission). Y-axis represents pitch; x-axis represents duration; the size/thickness of the circles/envelope represents volume; syllables or syllable boundaries are represented by the number of circles/found images or by the vertical lines in the envelope, depending on settings. One graphic represents the adult verbal model and the other graphic represents the child's imitation attempt.

During developmental play, play-based strategies and a mid-tech speech-generating device (i.e., GoTalk 20+^{TM}, Figure 2.2) were used to model and elicit productions within naturalistic interactions as a means to facilitate communicative competence (e.g., Girolametto et al., 1995, 1996; Robertson & Weismer, 1999; Romski & Sevcik, 2003). Example activities included social-interaction games (e.g., peek-a-boo, "I'm gonna get you"), communicative temptations (e.g., wind-up toys, bubbles, closed containers), singing songs, and sensory-motor activities (e.g., bean table and platform swing). Key treatment strategies included visual representations of the targets (e.g.,

a stuffed teddy bear to represent the target *teddy bear*) and spoken models of the targets in communicative contexts. The models were either spoken directly by the clinicians or modelled via the speech-generating device (SGD), which allowed programming of 25 different messages per overlay (see Figure 2.2). Each set of five targets was added to the device when they were introduced in treatment. Consistent with Weismer and Robertson (2006), clinicians aimed for at least 10 models of each of the five targets per session. In addition, clinicians utilized a combination of carrier phrases (e.g., "Here comes the… ?") and forced choices (e.g., "Do you want to color the *bunny* or the *butterfly?"*) as deemed beneficial to elicit a child's production of targets, as well as supplemental articulation and phonology cues as considered necessary to shape children's productions.



Figure 2.2. GoTalk 20+™ Communicator showcasing the overlay used for Pyrros during treatment phase 2. Images in the first row were included throughout the intervention, images in the second and third rows represent target sets 1 and 2 respectively.

In order to progress from one five-item target set to another, children had to a) master four of the five targets within a set during developmental play or b) participate in 12 treatment sessions for the respective target set. Mastery was defined as an accurate and spontaneous production of a multisyllabic target across two different treatment sessions; productions were considered spontaneous if at least three seconds had elapsed since the last verbal model, either produced by an adult or by the SGD. See supplementary material in DeThorne and colleagues (2015) for additional detail on intervention techniques used during motor practice and developmental play.

2.2.3.1.1 Treatment fidelity. Treatment sessions were coded, based on video review, by trained research assistants for the presence of key intervention components associated with motor practice and developmental play. Motor practice was considered to be implemented with fidelity if the clinician a) provided visual reference to syllable marking (e.g., tapped) for each target being treated in that session, and b) modeled each target a minimum of five times and/or if the child produced each target a minimum of five times. Based on these criteria, motor practice was completed with fidelity across 77% to 100% of participant's sessions (see Table 2.2). Developmental play was implemented with fidelity if the clinician provided at least 50 total models of targets per session, either spoken or through the speech-generating device. Based on this criterion, developmental play

was completed with fidelity across 94% to 100% of participant's available sessions (see Table 2.2). Inter-rater reliability based on 10% (13/125) of available sessions was 100% for both motor practice and developmental play.

Table 2.2. *Treatment Fidelity*

	Pyrros	Karis	Santiago	Angelo	Heidi	Conley
Motor Practice	97% (30/31)	77% (10/13)	86% (12/14)	100% (18/18)	88% (15/17)	88% (28/32)
Developmental Play	94% (29/31)	100% (13/13)	100% (14/14)	100% (18/18)	100% (17/17)	100% (32/32)

2.2.3.2 Assessment. After the initial assessment session conducted to determine eligibility, participants were assessed on all thirty targets (i.e., treatment and control) at five time points throughout the course of the study: prior to the start of the intervention (i.e., baseline), between target sets (i.e., assessments 2 and 3), post-treatment (i.e., assessment 4), and five weeks post-treatment (i.e., maintenance). Most children completed each assessment within one 45-minute session, with a maximum of 2 sessions per assessment. At each time point, two systematic assessment tasks were included to provide children the opportunity to produce the multisyllabic targets during two different activities: the object-play task and the card-labeling task (see Girolametto et al., 1995, 1996 for similar procedures). During the object-play task children were presented with a clear plastic box that included an object representing each of the 30 targets. Clinicians would offer a carrier phrase (e.g., "You want the ...") in an attempt to elicit a spontaneous production, and children were allowed to play with each object one at a time. During the card-labeling task, children were asked to verbally label picture-cards of each target one at a time (e.g., "What's this?"). During either assessment task, children were given 10 seconds per target to spontaneously produce the target. Assessment sessions were coded, based on video review, by trained research assistants. To examine reliability of outcome measures, research assistants independently coded all of the assessments (30/30). Based on point-by-point agreement, mean reliability for scoring the object-play task ranged from 97% to 99%; mean reliability for scoring the card-labeling task ranged from 97% to 100% (see Table 2.3). Consistent with DeThorne and colleagues (2015), percent accuracy of multisyllabic targets served as the dependent

variable; children were credited with production of a given target if it was accurately produced during either of the two tasks in order to maximize measurement stability.

Table 2.3. *Assessment Reliability*

	Pyrros	Karis	Santiago	Angelo	Heidi	Conley
Object-play	97% (93-100%)	99% (97-100%)	99% (97-100%)	99% (97-100%)	98% (90-100%)	98% (93-100%)
Card-labeling	98% (97-100%)	97% (90-100%)	98% (97-100%)	100% (100-100%)	99% (97-100%)	97% (93-100%)

2.3 Results

2.3.1 RQ1: Is there a functional relationship between the implementation of the speech-language intervention and the production of the multisyllabic speech targets? To address the first research question, the effectiveness of the intervention was examined through visual inspection via three forms of evidence: 1) within-participant comparison of progress in the treated vs. yet to be treated treatment targets consistent with single-case multiple-probe across behaviors experimental design; 2) within-participant comparison of production accuracy between the 15 treatment vs. the 15 control targets; and 3) within-participant comparison between pre- and post-intervention CDI scores.

2.3.1.1 Single-case experimental data. The primary form of evidence is data from the experimental single-case multiple-probe across behaviors design. This particular design evaluates intra-subject replications over time. Analyses centered on visual inspection of each child's production of the three target sets throughout the intervention. Each child's data is represented in a separate figure; see Figures 2.3 through 2.8. Each figure includes three vertically-stacked panels, one for each set of five speech targets. Note that the x-axis represents the session number and the y-axis represents the percent of the five targets produced correctly at each of the five assessment time points, or the percent of the five targets mastered during the development play section of treatment. The dotted line represents the introduction of treatment for each target set. As mentioned previously, correct production for any particular target was defined by accurate articulation of at least half the target phonemes in the correct order paired with verbal marking of all syllables. A combination of the card-labeling and object-play tasks were used at each of the five assessment points. Mastery during developmental play consisted of accurate and spontaneous production of a

target across two separate treatment sessions. Note that each assessment point presents a potential opportunity to demonstrate a treatment effect by vertical comparison of treated versus yet to be treated (i.e., untreated) targets. Consequently, each child presents data on an initial effect at assessment #2 and three opportunities for replication at assessments #3, #4, and maintenance.

2.3.1.1.1 Pyrros. Single-case data from Pyrros offered the strongest evidence for a positive intervention effect, see Figure 2.3. In the top panel, Pyrros' spontaneous and accurate productions of the first five-target set during the assessments increased from 0% during baseline to 20% at the close of the first treatment phase. During that same time frame, Pyrros' performance on the untreated target sets two and three remained unchanged at 0%. For the first target set, Pyrros did not achieve mastery of 4 of the 5 targets but progressed to the second five-item target set as he had participated in 12 treatment sessions. Considering a similar comparison for target set two (see middle panel), Pyrros' assessment performance increased from 0% to 40%, whereas his assessment performance for the third (yet to be treated) target set remained unchanged at 0% (see bottom panel). Once the third target set was treated, Pyrros' assessment performance on that set increased from 0% to 60%. Pyrros' production accuracy within target set either increased, for target sets 2 and 3, or was maintained, for target set 1, five-weeks post-treatment (i.e., maintenance assessment). Pyrros' assessment results suggest the multimodal speech-language intervention is functionally, or causally, related to the production of multisyllabic speech targets.

2.3.1.1.2 Karis. Similar to Pyrros, Karis' single-case multiple-probe across behaviors results suggest there is a functional relation between the intervention and the production of multisyllabic targets, albeit with a smaller magnitude (Figure 2.4). In the top panel, Karis' accurate productions of the first five-item target set increased from 0% during baseline to 40% at the close of the first treatment phase. In contrast, his performance on the untreated target sets two and three remained unchanged at 0%. In the middle panel, his performance on target-set two (middle-panel), increased from 0% to 20%, whereas his performance for the third untreated target set remained at 0%. Once his third target set was treated, his assessment performance increased from 0% to 20%. The five-week post-treatment maintenance assessment indicated Karis maintained his progress in target sets 1 and 3, and increased his production accuracy for target set 2.

2.3.1.1.3 Santiago. Santiago's first target set is indicative of an explicit functional relationship between intervention and accurate target production. Although replications associated with the

second target set indicate gains in one untreated target, visual analysis of production accuracy and slope in treated vs. untreated targets is still indicative of a functional relationship (see Figure 2.5). His first target set was associated with gains of 0% to 40%, whereas the untreated target sets (i.e. 2 and 3) remained at 0% during the same time period. Santiago's production of target set two (middle panel) increased from 0% to 80% once treated, whereas his performance for the third untreated target set demonstrated an increase in one of the five targets, moving from 0% to 20%. Although target set 3 did not remain at 0%, Santiago showed relative gains in the treated vs. the untreated targets. Once target set three was treated, his performance increased from 20% (at assessment 3) to 60%. Unfortunately, Santiago's performance on sets two and three treated targets decreased at future assessments, from 80% at assessment three to 40% at maintenance for target set 2, and from 60% at assessment four to 40% at maintenance for target set 3. In contrast, his performance on target set one increased from 40% at assessment two to 60% at maintenance.

2.3.1.1.4 Angelo. Although Angelo's gains in treated versus untreated targets are indicative of positive treatment effects, whether the relationship is causal in nature is less clear (see Figure 2.6). Specifically, the first target set was associated with gains of 0% to 80%. In contrast, his production of the untreated target set two remained at 0%, while his production of the untreated target set 3 increased from 0% to 20% during the same time period. Although target set 3 did not remain at 0%, Angelo showed relative gains in the treated versus the untreated targets. Angelo's production of the second set of targets increased from 0% to 40%, whereas his production accuracy for target set three, which had not yet been treated, increased from 20% (at assessment 2) to 40%. Once target set three was treated, it increased from 40% (at assessment 3) to 60% (post-treatment). The five-week post-treatment maintenance assessment indicated Angelo maintained his progress for all three target sets.

2.3.1.1.5 Heidi. Similar to Angelo, Heidi demonstrated gains in all target sets following treatment, but there was no clear evidence of experimental control across the target sets; see Figure 2.7. Specifically, in the top panel, Heidi's first target set increased from 0% to 40%. In contrast, the untreated target set 2 remained at 0%, while the untreated target set 3 increased from 0% to 20%. In the middle panel, her performance on target set two increased from 0% to 40% after treatment, whereas her performance for the third untreated target set increased from 20% to 60%. Once her third target set was treated, her assessment performance increased from 60% to 100%. Unfortunately, Heidi's performance on all three treated target sets decreased when comparing the

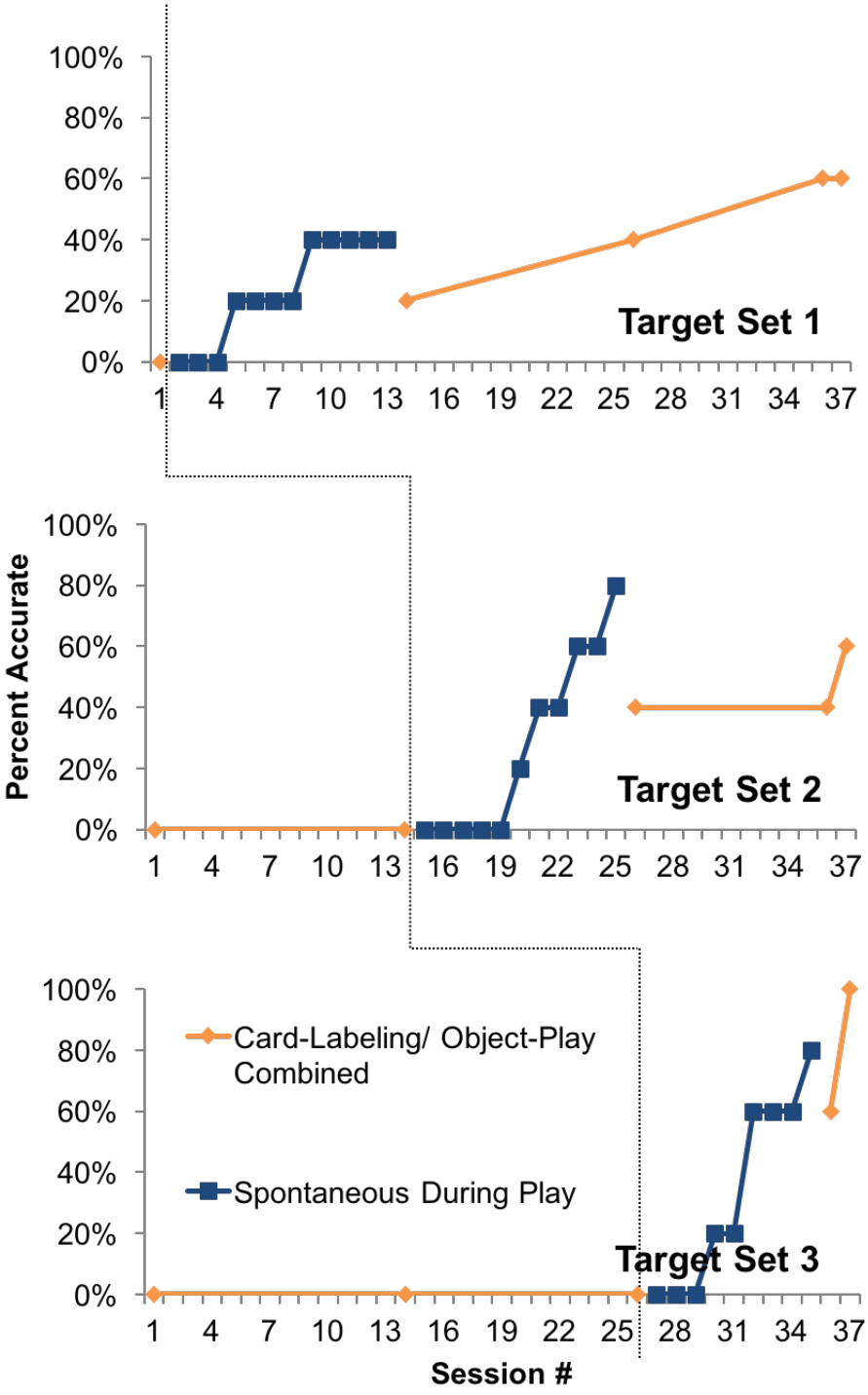


Figure 2.3. Pyrros combined object-play and card-labeling treatment and assessment data, with each panel representing one of the three different targets sets. Treatment data represents mastery of targets during developmental play (i.e., accurate and spontaneous production of a target during developmental play across two separate treatment sessions).

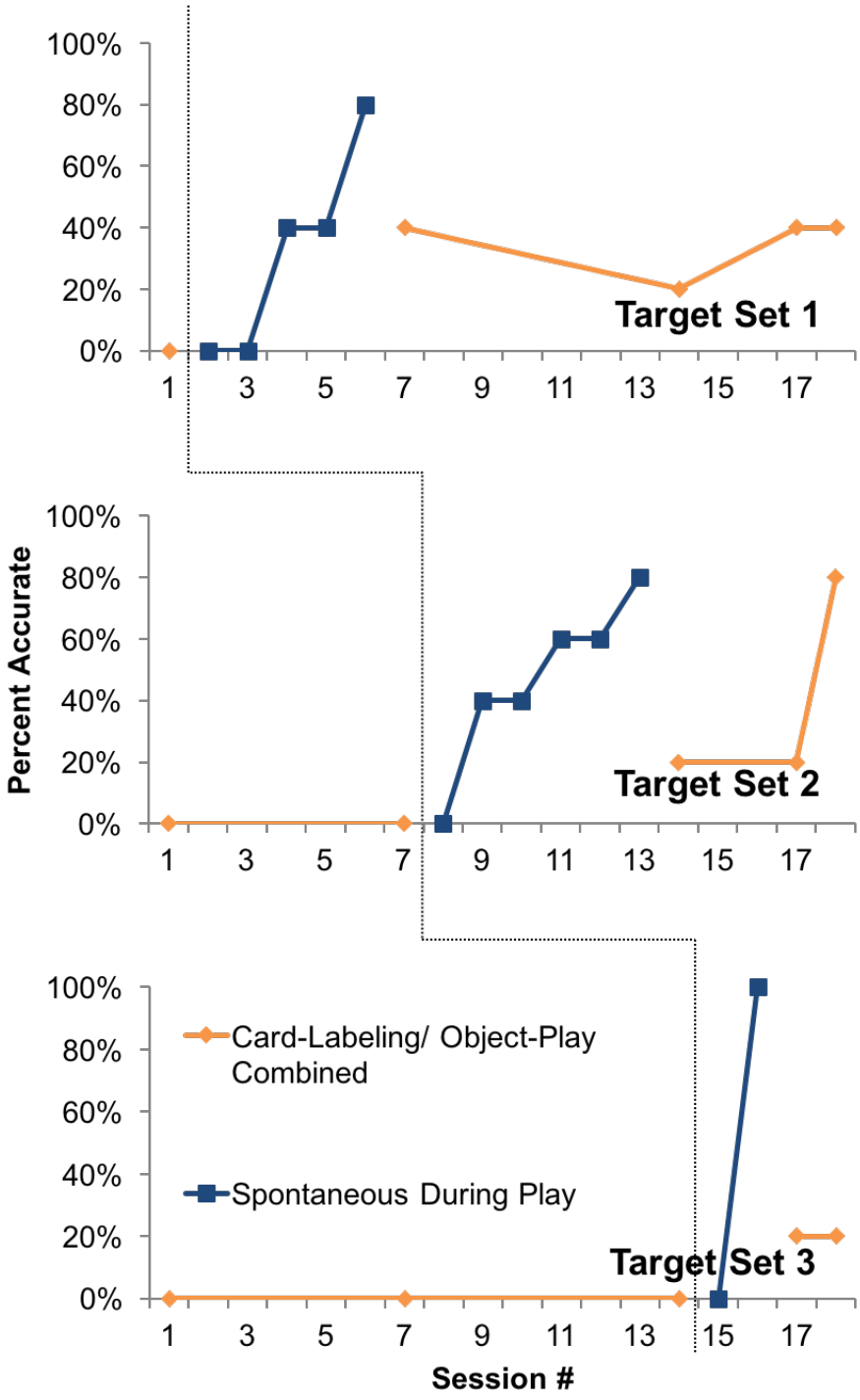


Figure 2.4. Karis combined object-play and card-labeling treatment and assessment data, with each panel representing one of the three different targets sets. Treatment data represents mastery of targets during developmental play (i.e., accurate and spontaneous production of a target during developmental play across two separate treatment sessions).

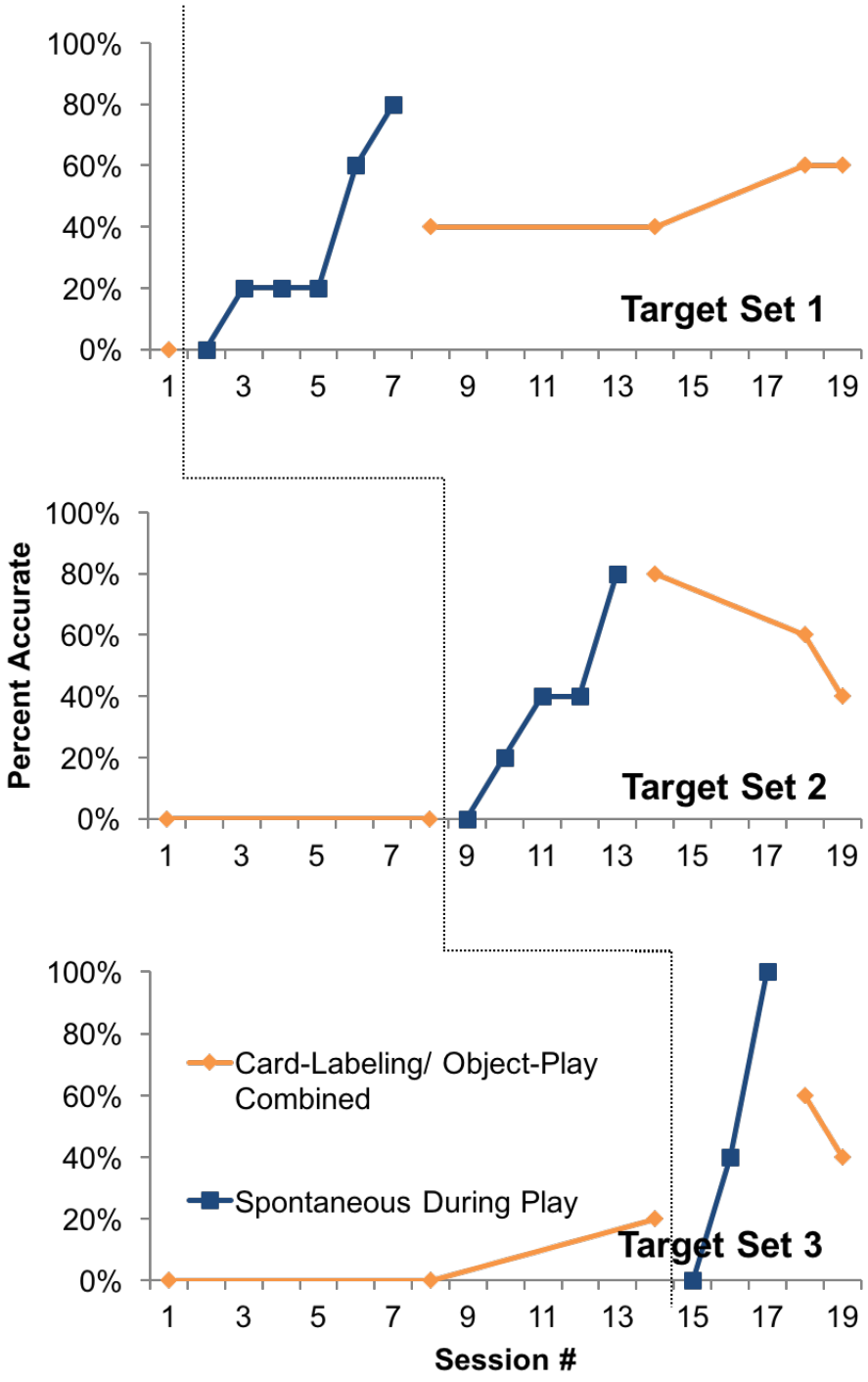


Figure 2.5. Santiago combined object-play and card-labeling treatment and assessment data, with each panel representing one of the three different targets sets. Treatment data represents mastery of targets during developmental play (i.e., accurate and spontaneous production of a target during developmental play across two separate treatment sessions).

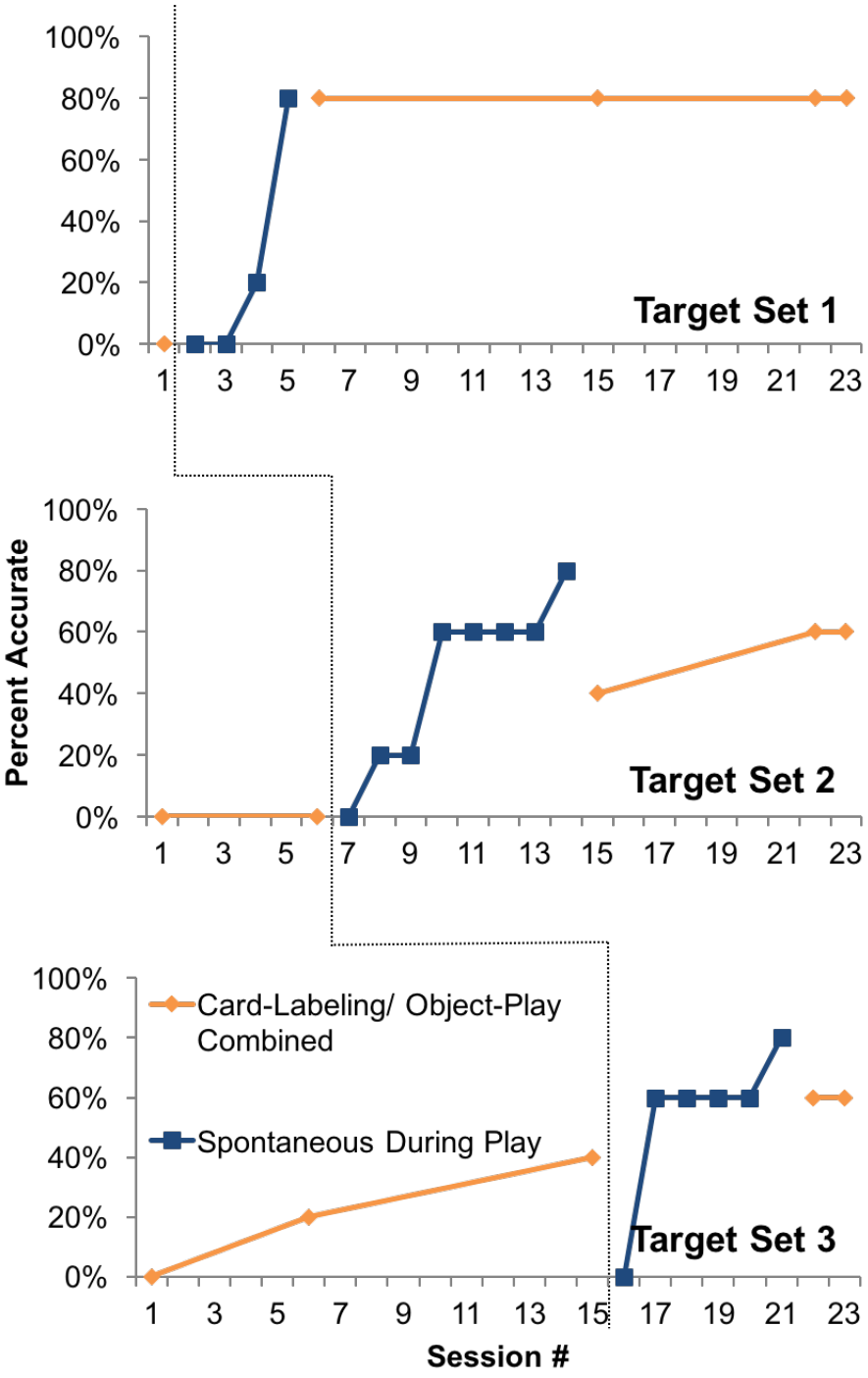


Figure 2.6. Angelo combined object-play and card-labeling treatment and assessment data, with each panel representing one of the three different targets sets. Treatment data represents mastery of targets during developmental play (i.e., accurate and spontaneous production of a target during developmental play across two separate treatment sessions).

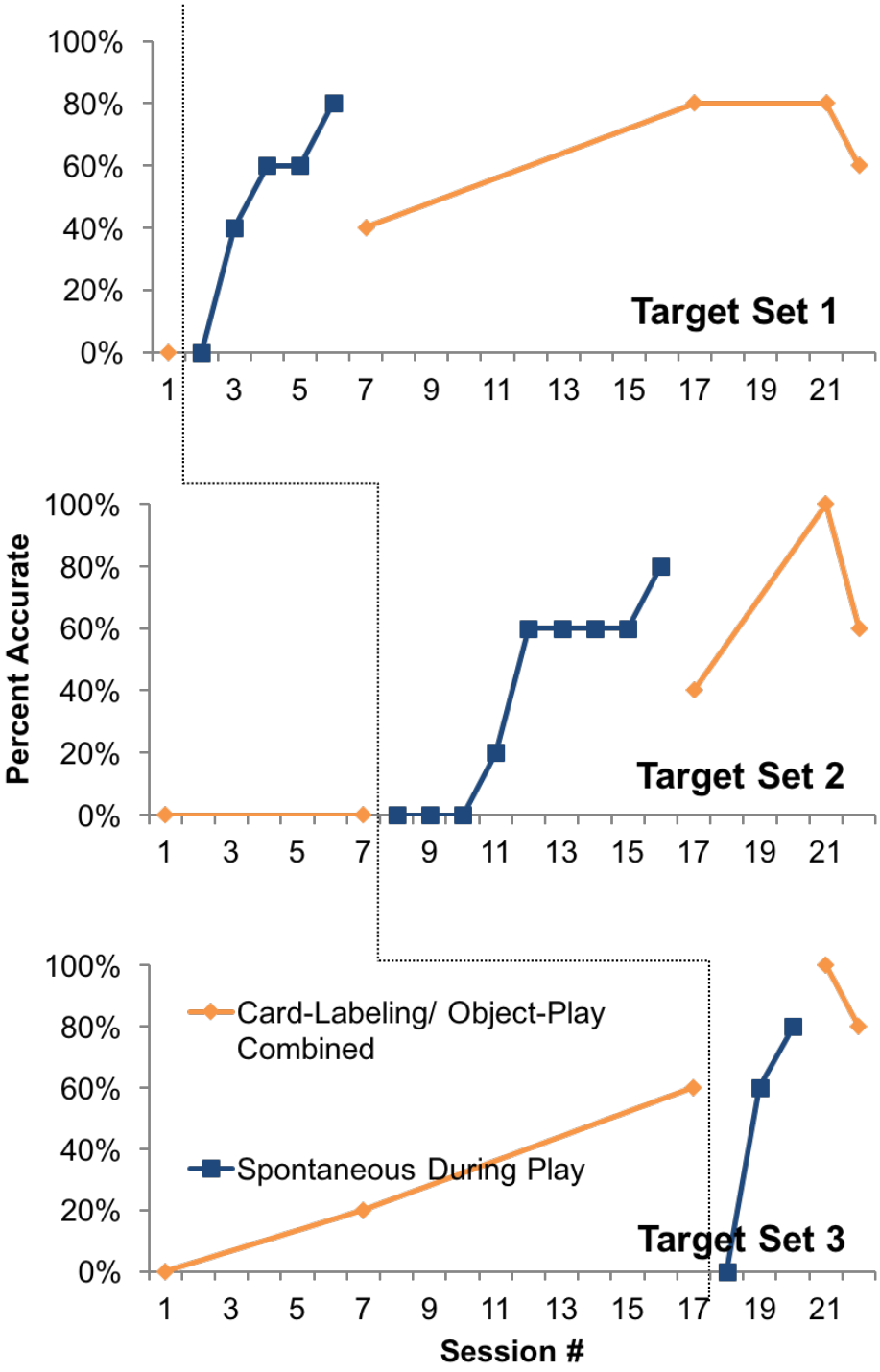


Figure 2.7. Heidi combined object-play and card-labeling treatment and assessment data, with each panel representing one of the three different targets sets. Treatment data represents mastery of targets during developmental play (i.e., accurate and spontaneous production of a target during developmental play across two separate treatment sessions).

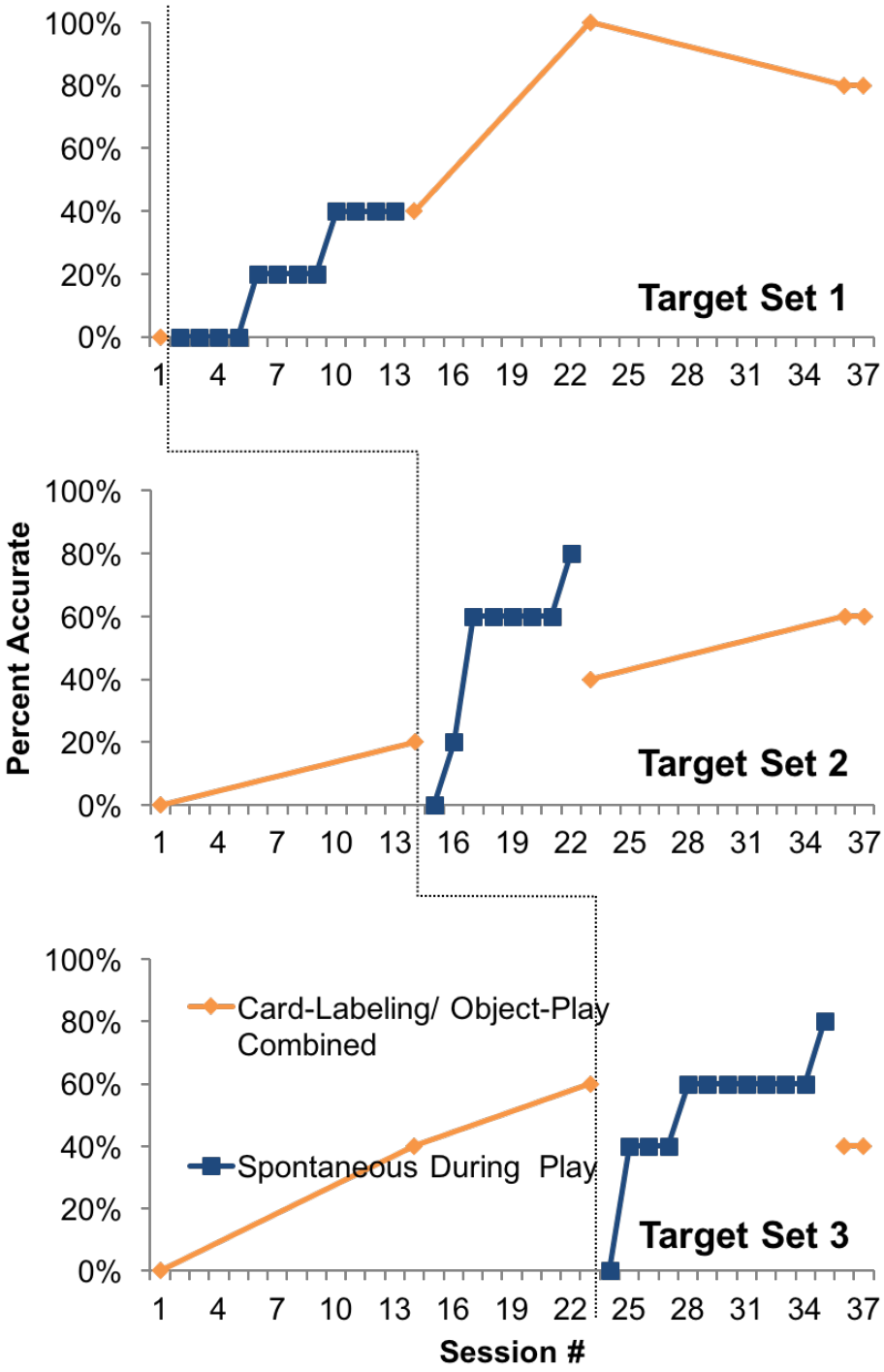


Figure 2.8. Conley combined object-play and card-labeling treatment and assessment data, with each panel representing one of the three different targets sets. Treatment data represents mastery of targets during developmental play (i.e., accurate and spontaneous production of a target during developmental play across two separate treatment sessions).

post-treatment performance with the maintenance performance, with a 20% to 40% decrease (i.e., one to two targets).

2.3.1.1.6 Conley. Conley demonstrated gains in target sets one and two following treatment but did not demonstrate gains following treatment for target set three (see Figure 2.8). In the top panel, Conley's first target set increased from 0% to 40%. During the same time period, the untreated target set 2 increased to 20% and the untreated target set 3 increased to 40%. For target set 1, Conley did not achieve mastery of 4 of the 5 targets but progressed to the second five-item target set as he had participated in 12 treatment sessions. In the middle panel, his performance on target set 2 increased from 20% to 40%, and his performance for the third untreated target set increased from 40% to 60%. Moreover, Conley's performance on treated target set three decreased from assessment three at 60% to 40% at post-treatment and maintenance. When comparing post-treatment to maintenance performance, Conley was able to maintain his progress five weeks after the end of treatment. Overall, Conley's assessment results do not provide any clear evidence of experimental effects.

***

In sum, single-case experimental data indicated a clear functional relationship between the intervention and increased target production in half of the participants (i.e., Pyrros, Karis, & Santiago). Although the other half of the participants demonstrated gains in targets, they did not demonstrate a replicated treatment effect.

2.3.1.2 Within-participant comparison of treatment versus control targets. Though not inherently experimental in nature, the within-participant comparison between treatment versus control targets at each assessment point provides additional evidence related to potential treatment effects. Recall that treatment and control targets were matched in terms of syllable number, sematic category, and phonetic complexity. Consequently, if participants perform better on treatment relative to control targets it provides supporting evidence of a positive treatment effect. Note from Table 2.4 that other than baseline, at which point participants scored 0% across both treatment and control targets, participants scored higher on the treatment targets relative to control targets on all but two of the twenty-four assessment points. The only exceptions were Santiago and Pyrros at Assessment #2. Santiago scored lower on treatment vs. control targets, correctly producing 2/15 treatment targets and 3/15 control targets; Pyrros scored equally well on the treatment and control targets, correctly producing 1/15 targets from each type (treatment v. control). Specifically at the

Table 2.4. *Percent accuracy throughout the five assessment points for the 15 treatment (Tx) and the 15 control (Cntrl) targets.*

	Pyrros		Karis		Santiago		Angelo		Heidi		Conley	
	Tx	Cntrl	Tx	Cntrl	Tx	Cntrl	Tx	Cntrl	Tx	Cntrl	Tx	Cntrl
Baseline	0	0	0	0	0	0	0	0	0	0	0	0
Assessment 2	7	7	13	0	13	20	33	7	20	7	33	0
Assessment 3	27	13	13	0	47	27	53	20	60	40	67	27
Post-treatment	53	20	27	0	47	13	67	33	93	60	60	27
Maintenance	73	47	47	20	47	27	67	60	67	47	60	47

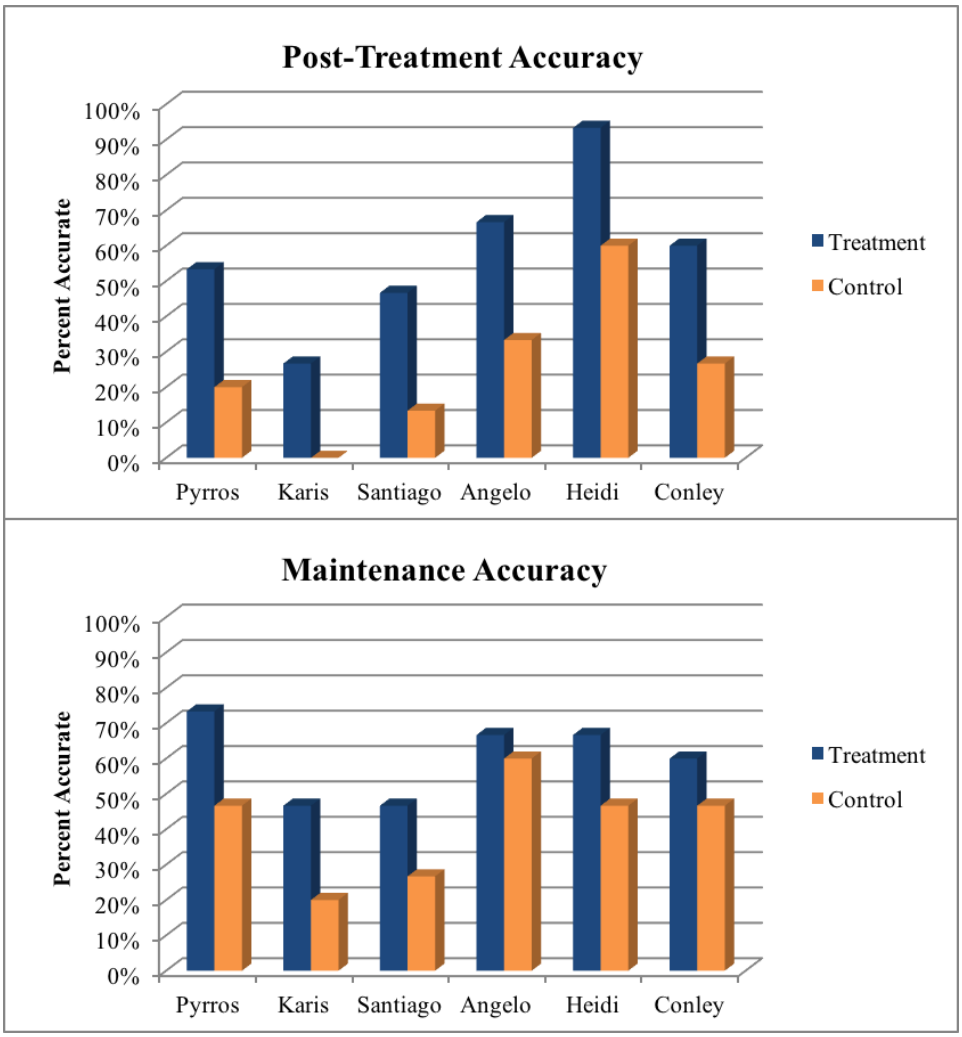


Figure 2.9. Post-treatment (top panel) and maintenance (bottom panel) percent production accuracy of the 15 treatment vs. the 15 control targets.

post-treatment assessment, all six children performed better on treatment targets relative to control targets (see Figure 2.9, top panel). Maintenance results indicated all children except Heidi were able to maintain or increase their treatment and control gains five-weeks post-treatment, with correct production of 7 to 11 treatment targets versus 3 to 9 control targets. In contrast, Heidi's production of both treatment and control targets decreased from post-treatment to maintenance, by 4 and 2 words respectively, with increased production of treatment vs. control targets [5] (see Figure 2.9, bottom panel).

2.3.1.3 Within-participant comparison between pre- and post-intervention CDI scores. A third and final form of evidence related to potential treatment effects is presented in Table 2.5: each child's pre- and post-intervention communicative development expressive vocabulary inventories. Results demonstrated all participants with available data increased their expressive vocabulary and more specifically, their expressive multisyllabic vocabulary inventories during the course of intervention. Representing the range of gains, Pyrros' mother reported a gain of 29 multisyllabic words from the CDI during a course of 33 weeks, whereas Conley's mother reported a gain of 129 words from the same inventory during the course of 37 weeks between pre- and post-CDI administration. Although these data do not provide experimental control for confounding effects such as maturation, they do provide an additional form of converging evidence.

Table 2.5. *Pre- and Post-Intervention CDI Scores*

	Pyrros		Karis		Santiago		Angelo		Heidi		Conley	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
CDI-expressive all	82	198	276	NA	183	253	154	253	88	177	64	393
CDI-expressive multisyllabic	22	51	100	NA	61	93	52	96	41	79	21	150

[5] The within-participant comparison of treatment versus control targets in the present chapter further extends the within-subject experimental analysis in DeThorne and colleagues (2015) by including object-play data in addition to card-labeling data. To facilitate comparison with results published in 2015, within-subject experimental group outcomes are also presented (i.e., means and statistical analyses using two paired-sample *t*-tests). Results indicated children were able to produce approximately 58% (Range = 27-93%) of the treated targets on demand using spoken language, compared to only 26% (Range = 0-60%) of the control targets by the end of the intervention ($p < .001$, $d = 11.84$). Similar differences between the treated targets ($M = 60\%$, Range = 47-73%) and the control targets ($M = 41\%$, Range = 20 -60%) remained at the five-week post-treatment maintenance session ($p = .002$, $d = 2.42$).

2.3.2 RQ2: What information does the within-participant data provide relative to the experimental group results published in DeThorne and colleagues (2015)? To address the second research question, we highlight three contributions, not mutually exclusive, provided by the within-participant data relative to the experimental group results published in DeThorne and colleagues (2015). First, within-participant data provides a more in-depth characterization of participant performance across time by highlighting the rate in which each child achieved mastery of targets and the rate of acquisition during developmental play. In regards to rate of mastery during developmental play, single-case results indicated that all children except Pyrros were able to master[6] 1-5 targets in at least two of the three target sets after only two treatment sessions (see Figures 2.3 - 2.8). While it took Pyrros at least four treatment sessions to master the first target within a set (see top and bottom panel of Figure 2.3), Karis was able to master all five targets of target set 3 after only two treatment sessions (see bottom panel of Figure 2.4). Overall, the majority of children were able to master at least four of the five targets per set within 12 treatment sessions (Range = 67 – 100%); see Figures 2.3 - 2.8. Pyrros and Conley were the only two children that did not progress to the next target set by achieving mastery, but this was only the case for the first target set (top panel of Figures 2.3 and 2.8). Although Pyrros only mastered 2/5 targets during treatment phase 1, he was able to spontaneously and accurately produce 3/5 targets during developmental play at least once throughout 12 treatment sessions. Similarly, even though Conley only mastered 2/5 targets during treatment phase 1, he was able to produce 4/5 targets at least once during developmental play throughout 12 treatment sessions. In regards to rate of acquisition during developmental play, all children except Pyrros were able to accurately and spontaneously produce at least 1/15 targets after only one treatment session (Range = 5 to 9). After two treatment sessions, all children, including Pyrros, were able to accurately and spontaneously produce at least 4/15 targets during developmental play (Range = 4 to 8); see Table 2.6. Across target sets and treatment phases, all children were able to spontaneously and accurately produce a majority of the treated targets at least once during developmental play using spoken language (Range = 80-100%) (data not shown). Although not experimental in nature, these findings are notable given that children were not able to produce any of the targets at the beginning of the study.

[6] i.e., spontaneous and accurate production of a target across two separate treatment sessions during developmental play.

Table 2.6. *Rate of Acquisition of Targets During Developmental Play for the First Two Treatment (Tx) Sessions per Target Set*

	Pyrros		Karis		Santiago		Angelo		Heidi		Conley	
	Tx 1	Tx 2	Tx 1	Tx 2	Tx 1	Tx 2	Tx 1	Tx 2	Tx 1	Tx 2	Tx 1	Tx 2
Target Set 1	NA	Night night	Kitty, Person	Shovel	Chicken, Penguin	Chicken	Sleepy	Tooth-brush	Party, Pillow, Drawer	Party, Pillow, Coffee	NA	NA
Target Set 2	NA	Drink it, Firetruck	Yucky, Donkey	Yucky, Donkey	Pony, Raisin	Pony, Garden, Spaghetti	Happy	Happy, Radio, Drawer	NA	NA	Puppy, Finger	Puppy, Zipper
Target Set 3	NA	Teddy bear	Bunny, Penny, Drawer, Gentle, Rocking chair	Bunny, Penny, Drawer, Gentle, Rocking chair	Dirty	Dirty, Oven, Giraffe, Vacuum	Party, Tummy, Jacket	Party, Tummy, Jacket, Candy	Hello, Donkey, Button	Hello, Donkey, Button, Empty, High chair	Apple, Bunny, Window	Apple, Bunny

A second contribution of the within-participant data is that it provides assessment data across multiple time points, in this case across five assessment points throughout the course of the intervention (see Table 2.4). Results indicated that from baseline to post-treatment (i.e., end of treatment sessions) five of the six children were able to increase and/or maintain their production of treatment targets as the intervention progressed. More specifically, three of the six participants (i.e., Pyrros, Angelo, and Heidi) increased their production of treatment targets after every treatment phase (i.e., at assessment 2, assessment 3, and post-treatment); this was also the case for their control targets, albeit with a smaller magnitude. For the other two children who were able to increase and/or maintain their production of treatment targets, Karis did not make any treatment (or control) gains after the second treatment phase (i.e., from assessment 2 to assessment 3). Although Santiago did not make any gains after the third and last treatment phase (i.e., from assessment 3 to post-treatment), he was able to maintain his gains on treatment targets (7/15 targets), whereas his production accuracy for control targets decreased from 4/15 targets at assessment 3 to 2/15 targets at post-treatment. Conley was the only participant whose production of treatment targets decreased while the speech-language therapy was still ongoing; he produced 10/15 targets at assessment 3 versus 9/15 targets at post-treatment. Moreover, in regards to production accuracy from post-treatment to the five-week post-treatment maintenance session, all children except Heidi were able to increase or maintain their production accuracy, and production of treated targets was greater than that of control targets for all children at these last two time points.

A third contribution of within-participant data is that it highlights the individual variability often masked by group outcomes. Experimental group findings in DeThorne and colleagues (2015) demonstrated a significant difference between VocSyl and the active Control group at the maintenance time point, with increased target accuracy for the VocSyl group. However, the single-case experimental design revealed more conservative estimates of treatment effect at the individual level. Specifically, the design revealed an explicit functional relationship between the intervention and outcome data for only 3 of the 6 children (i.e., Pyrros, Karis, & Santiago). Even for the three participants in which we were able to establish a causal relationship, the magnitude of the effect differed by participant. Whereas Pyrros demonstrated the strongest effect, such effect was attenuated for Karis, and in the case of Santiago we observe a decrease in production accuracy over time for two of his five-item target sets (Figure 2.5). Similarly, even though DeThorne and colleagues (2015) reported mean differences between children's gains on the treated vs. control targets both at post-treatment and at the maintenance time points for the VocSyl condition, the within-participant comparison of treated versus control targets in the present study offered information about both the consistency and variability of this trend across individual participants. Whereas Karis did not acquire any of the control targets from baseline to post-treatment while acquiring four of the treated targets, Heidi was able to acquire nine of the control targets while acquiring 14 of the 15 treated targets post-treatment.

2.4. Discussion

In sum, this study a) provides additional evidence for the use of a multimodal, integrated speech-language intervention to facilitate multisyllabic productions, b) illustrates the potential role of technology in facilitating speech-language development, and c) highlights single-case design as a complementary method to experimental group design (see DeThorne et al., 2015 for group results). In regard to treatment effects, converging forms of within-participant evidence provided support for a positive treatment effect for all children, although with substantial individual variability. Overall, the use of a multimodal, integrated speech-language approach facilitated the production of new vocabulary as well as the use of a relatively new phonological form (i.e., syllable combination). Nevertheless, single-case experimental data offered evidence that the intervention was causally related to the production of multisyllabic targets for only half of our participants (3/6); confounding influences such as practice, maturation, and generalization could

not be ruled out as contributing factors to the progress of the other three children. In the remainder of the discussion I will highlight the role of technology within intervention and the strengths and limitations of single-case methodology, and of focusing on individual outcomes more broadly.

During treatment, the present study incorporated a novel form of technology, VocSyl, in addition to a mid-tech speech-generating device, the GoTalk 20+, thereby highlighting the role of technology in speech-language development. VocSyl offered an opportunity to provide real-time visualizations to children in regard to number of syllables, pitch, rate, and volume. The strong appeal of electronic technology for children paired with the dynamic nature of VocSyl was motivating for our participants, and likely influenced the higher motor practice treatment fidelity for the VocSyl participants (see Table 2.2) relative to the Pacing Board participants as reported in DeThorne and colleagues (2015). The use of VocSyl, however, was not without its limitations. For example, the syllable algorithm was not always accurate within applied contexts (see DeThorne et al., 2015 for a more thorough discussion of the limitations of the software). In addition, the GoTalk offered a second form of technology. It was integrated throughout intervention to provide verbal models and visual support of individual targets, and offered each child an alternative to spoken language. Although the study was not designed to examine individual treatment components, the results add to a growing literature that supports the use of technology to facilitate communication, both as a form of visual feedback (Aparicio Betancourt et al., 2017; Bernard-Opitz, Sriram, & Sapuan, 1999; Hailpern, Karahalios, & Halle, 2009) and as a form of augmentative and alternative communication (e.g., Romski et al., 2010).

A major strength of single-case designs is the ability to study individual outcomes, while maintaining the ability to demonstrate experimental control. As stated by Kazdin (2011), "by studying the individual, the experimenter could see lawful behavioral processes that might be hidden in averaging performance across several subjects, as is commonly done in group research" (p. 14). The emphasis on individual outcomes provides the opportunity to focus on vital information about the uniqueness of the individual. This is seen, for example, in the case of Conley, in which his data contribute to the positive treatment effects observed at the group level, whereas individually, the positive treatment effects do not provide evidence of a causal relationship (Figure 2.8), and instead appear to be related to a substantial language spurt that may have been positively influenced by treatment (e.g., see Table 2.4 in which production of treatment targets was greater than control targets across time points). The richness of individual data unmasks individual

variability, and thereby offers an opportunity to better match individual children's profiles with specific interventions. For example, it allows you to explore questions such as, who would benefit the most from this type of intervention? And what are some of the characteristics associated with a stronger treatment effect? This level of information allows educators and clinicians to more directly relate the evidence-based interventions with the students or patients they are working with.

Within-participant methodology such as single-case designs also increases the feasibility for clinicians and educators to implement the design relative to group designs. Although it is generally not feasible for clinicians and educators to conduct well-controlled between-group experimental studies, the data collection procedures associated with single-case designs are very similar to that required for documenting extent of progress toward treatment goals in educational and medical settings. With the careful addition of a control condition (e.g., comparable untreated behaviors, settings) clinicians and educators can experimentally assess the intervention effect while controlling for confounding influences such as maturation. Such data could be particularly powerful for making decisions about intervention and/or educational opportunities, and in trying to persuade potential funding sources such as insurance companies.

The nature of single-case designs also encourages an educator/clinician to tailor the intervention to the individual, thereby allowing for more variation in procedures in comparison with group designs. During our intervention, the use of specific treatment strategies was influenced by the child's individual response. To illustrate, Pyrros was particularly responsive to simplifying the phonological complexity of the target (e.g., /taɪ͡taɪ͡/ followed by /taɪ͡gaɪ͡/ and finally /taɪ͡gɚ/ for *tiger*), and pictorial representation of the target (e.g., drawings of an eye and a truck, followed by drawings of a fire and a truck for *firetruck*), so these strategies were used more frequently. Overall, single-case research is relevant for defining clinical and educational practices at the level of the individual learner given the emphasis of fields such as speech-language pathology and special education is on the individual patient or student. Although single-case designs are well integrated within the field of special education, they have yet to be fully integrated within the field of communication sciences and disorders.

One of the most challenging limitations for the application of single-case methodology to speech and language development, is that learned skills cannot be unlearned, thereby rendering some single-case designs, such as reversal designs, untenable. Reversal or ABAB designs alternate the baseline (A phase) and the intervention (B phase) to evaluate the effect of the intervention.

However, when a child learns a new sound or word, they are not going to unlearn it because the condition changes; that is, the level of performance will not necessarily decrease after the treatment is discontinued. Although other single-case designs, such as multiple baseline across participants and behaviors, are more amenable to developmental skills, another challenge is that generalization of skills often leads to loss of experimental control. In the present study, experimental control associated with the multiple probe across behaviors design was contingent on the idea that learning the targeted multisyllabic productions would not elicit general learning of other multisyllabic productions, such as the yet to be treated targets. However, from a developmental standpoint, the most powerful interventions will trigger generalization. To illustrate, experimental control was not established for Angelo because he began to produce untreated targets together with treated targets. Specifically, his production of one untreated target at assessment 2 (i.e., *party*), and two untreated targets at assessment 3 (i.e., *party* and *tummy*) jeopardized experimental control, even though this may actually have been evidence of generalized treatment effects (see Figure 2.6). For example, one of his first treatment targets was the word *dirty* (/dɝti/ or /dɝɾi/) which shared very similar phonological features with the untreated targets *party* (/pɑ͡ɹɹi/ or / pɑ͡ɹti/) and *tummy* (/tʌmi/). Angelo first produced the target *dirty* during developmental play of treatment phase 1, and subsequently produced the targets *party* and *tummy*. Consequently, it is possible that treating the word *dirty* helped facilitate his acquisition of the untreated targets *party* and *tummy* and thereby jeopardized the very experimental control needed to demonstrate a causal treatment effect.

In sum, "information from groups and information from individuals contribute separate but uniquely important sources of information" to applied research (Kazdin, 2011, p. 12). In the present study we found that single-case methodology offered useful converging evidence for the treatment effects reported in DeThorne et al. (2015) but also uncovered important forms of individual variability that could be used to inform educational and clinical practice. It is interesting to note that even though single-case research is sometimes presented as a less rigorous methodology (e.g., ASHA, n.d.), in the present project the single-case design yielded a more conservative experimental outcome than did the group design. Moving forward, our hope is to increase awareness and acceptance of the strengths associated with science focused on understanding individual learning trajectories, and for patient-centered organizations such as the American Speech-Language-Hearing Association to explicitly recognize single-case designs in

their levels of evidence (see ASHA, n.d.) as another method that can be used to establish evidence-based interventions (cf. Justice & Fey, 2004).

Acknowledgments: This project was funded by Autism Speaks (#5744) and the University of Illinois Research Board; MAB was additionally funded by the Neuroscience Program Fellowship, the Goldstick Family Doctoral Fellowship for the study of communication disorders, and the Child Language & Literacy Laboratory. We are appreciative of all participating families; Josh Hailpern and Reed La Botz for assistance with VocSyl; as well as assistance from the Child Language & Literacy Lab, particularly: Ellen Bogue, Kristin Lyons, Lauren Ragins, Monique Kammo, Johanna Drucker, Emily Zimmerman, Kori Zorina, Meghan Schassler, and Kathy Dejmek.

References

American Speech-Language-Hearing Association. (n.d.). Steps in the process of evidence-based practice. Retrieved July 28, 2017 from http://www.asha.org/Research/EBP/Assessing-the-Evidence/

Aparicio Betancourt, M., Dethorne, L. S., Karahalios, K., & Kim, J. G. (2017). Skin conductance as an *in situ* marker for emotional arousal in children with neurodevelopmental communication impairments: Methodological considerations and clinical implications. *ACM Transactions on Accessible Computing (TACCESS)*, *9*(3), 8. doi: 10.1145/3035536

Bernard-Opitz, V., Sriram, N., & Sapuan, S. (1999). Enhancing vocal imitations in children with autism using the IBM speech viewer. *Autism, 3*(2), 131–147. https://doi.org/10.1177/1362361399003002004

Binger, C., & Light, J. (2007). The effect of aided AAC modeling on the expression of multi-symbol messages by preschoolers who use AAC. *Augmentative and Alternative Communication*, *23*(1), 30-43. http://dx.doi.org/10.1080/07434610600807470

DeThorne, L. S., Johnson, C. J., Walder, L., & Mahurin-Smith, J. (2009). When "Simon says" doesn't work: Alternatives to imitation for facilitating early speech development. *American Journal of Speech-Language Pathology*, *18*(2), 133-145. doi:10.1044/1058-0360(2008/07-0090)

DeThorne, L., Aparicio Betancourt, M., Karahalios, K., Halle, J., & Bogue, E. (2015). Visualizing syllables: Real-time computerized feedback within a speech–language intervention. *Journal of Autism and Developmental Disorders*, *45*(11), 3756-3763. doi:10.1007/s10803-014-2274-8

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates communicative development inventories* (2nd ed.). Baltimore, MD, USA: Paul H. Brookes.

Girolametto, L., Pearce, P. S., & Weitzman, E. (1995). The effects of focused stimulation for promoting vocabulary in young children with delays: A pilot study. *Journal of Children's Communication Development*, *17*(2), 39-49. https://doi.org/10.1177/152574019501700205

Girolametto, L., Pearce, P. S., & Weitzman, E. (1996). Interactive focused stimulation for toddlers with expressive vocabulary delays. *Journal of Speech, Language, and Hearing Research*, *39*(6), 1274-1283. doi:10.1044/jshr.3906.1274

Hailpern, J., Harris, A., La Botz, R., Birman, B., & Karahalios, K. (2012). Designing visualizations to facilitate multisyllabic speech with children with autism and speech delays. *Proceedings of the ACM Designing Interactive Systems (DIS) Conference, Newcastle upon Tyne, UK,* 126-135. doi: 10.1145/2317956.2317977

Hailpern, J., Karahalios, K., & Halle, J. (2009, April). Creating a spoken impact: Encouraging vocalization through audio visual feedback in children with ASD. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp.453–462). ACM.

Hailpern, J., Karahalios, K., DeThorne, L., & Halle, J. (2010). VocSyl: Visualizing syllable production for children with ASD and speech delays. *Proceedings of the 12th international ACM SIGACCESS conference on Computers and accessibility, Orlando, FL, USA,* 297-298. doi: 10.1145/1878803.1878879

Harjusola-Webb, S. M., & Robbins, S. H. (2012). The effects of teacher-implemented naturalistic intervention on the communication of preschoolers with autism. *Topics in Early Childhood Special Education*, *32*(2), 99-110. doi: 10.1177/0271121410397060

Hayden, D., & Square, P. (1999). *Verbal motor production assessment for children*. San Antonio, TX, USA: The Psychological Corporation.

Highman, C., Hennessey, N., Sherwood, M., & Leitão, S. (2008). Retrospective parent report of early vocal behaviours in children with suspected childhood apraxia of speech (sCAS). *Child Language Teaching and Therapy, 24*(3), 285-306. https://doi.org/10.1177/0265659008096294

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional children*, *71*(2), 165-179. https://doi.org/10.1177/001440290507100203

Jakielski, K. J., Maytasse, R., & Doyle, E. (2006, November). *Acquisition of phonetic complexity in children 12–36 months of age.* Poster presented at the annual convention of the American Speech–Language–Hearing Association, Miami, FL, USA.

Johnson, C.J. (2006). Getting started in evidence-based practice for childhood speech-language disorders. *American Journal of Speech-Language Pathology, 15*(1), 20-35. doi:10.1044/1058-0360(2006/004)

Justice, L. M., & Fey, M. E. (2004). Evidence-based practice in schools: Integrating craft and theory with science and data. *The ASHA Leader*, *9*(17), 4-32. doi: 10.1044/leader.FTR1.09172004.4

Kazdin, A. E. (2011). *Single-case research designs: Methods for clinical and applied settings* (2nd ed.). New York, NY, USA: Oxford University Press.

Kumin, L., Councill, C., & Goodman, M. (1995). The pacing board: A technique to assist the transition from single word to multiword utterances. *Infant-Toddler Intervention: The Transdisciplinary Journal, 5*(1), 23-30.

lacono, T., Mirenda, P., & Beukelman, D. (1993). Comparison of unimodal and multimodal AAC techniques for children with intellectual disabilities. *Augmentative and Alternative Communication*, *9*(2), 83-94. http://dx.doi.org/10.1080/07434619312331276471

lacono, T., Mirenda, P., & Beukelman, D. (1993). Comparison of unimodal and multimodal AAC techniques for children with intellectual disabilities. *Augmentative and Alternative Communication*, *9*(2), 83-94. http://dx.doi.org/10.1080/07434619312331276471

Law, J., Boyle, J., Harris, F., Harkness, A., & Nye, C. (2000). Prevalence and natural history of primary speech and language delay: Findings from a systematic review of the literature. *International Journal of Language and Communication Disorders*, *35*(2), 165-188.

Morris, S. R. (2009). Test–retest reliability of independent measures of phonology in the assessment of toddlers' speech. *Language, Speech, and Hearing Services in Schools*, *40*(1), 46-52. doi:10.1044/0161-1461(2008/07-0082)

Preston, J. L., & Edwards, M. L. (2007). Phonological processing skills of adolescents with residual speech sound errors. *Language, Speech, and Hearing Services in Schools*, *38*(4), 297-308. doi:10.1044/0161-1461(2007/032)

Robertson, S. B., & Weismer, S. E. (1999). Effects of treatment on linguistic and social skills in toddlers with delayed language development. *Journal of Speech, Language, and Hearing Research*, *42*(5), 1234-1248. doi:10.1044/jslhr.4205.1234

Romski, M. A., & Sevcik, R. A. (2003). Augmented input: Enhancing communication development. In J. Light, D. R. Beukelman, & J. Reichle (Eds.), *Communicative competence for individuals who use AAC* (pp. 147-162). Baltimore, MD, USA: Paul H. Brookes.

Romski, M., Sevcik, R. A., Adamson, L. B., Cheslock, M., Smith, A., Barker, R. M., & Bakeman, R. (2010). Randomized comparison of augmented and nonaugmented language interventions for toddlers with developmental delays and their parents. *Journal of Speech, Language, and Hearing Research*, *53*(2), 350-364. doi:10.1044/1092-4388(2009/08-0156)

Rose, T. (2016). *The end of average: How we succeed in a world that values sameness*. New York, NY, USA: HarperCollins.

Schlosser, R. W., & Wendt, O. (2008). Effects of augmentative and alternative communication intervention on speech production in children with autism: A systematic review. *American Journal of Speech-Language Pathology*, *17*(3), 212-230. doi:10.1044/1058-0360(2008/021)

Squires, J., & Bricker, D. (2009). *Ages & stages questionnaires: A parent-completed child monitoring system* (3rd ed.). Baltimore, MD, USA: Paul H. Brookes.

Strand, E. A., & Debertine, P. (2000). The efficacy of integral stimulation intervention with developmental apraxia of speech. *Journal of Medical Speech-Language Pathology*, *8*(4), 295-300.

Strand, E. A., Stoeckel, R., & Baas, B. (2006). Treatment of severe childhood apraxia of speech: A treatment efficacy study. *Journal of Medical Speech-Language Pathology*, *14*(4), 297-307.

Tager-Flusberg, H., Rogers, S., Cooper, J., Landa, R., Lord, C., Paul, R., ... Yoder, P. (2009). Defining spoken language benchmarks and selecting measures of expressive language development for young children with autism spectrum disorders. *Journal of Speech, Language, and Hearing Research*, *52*(3), 643-652. doi:10.1044/1092-4388(2009/08-0136)

Velleman, S. L. (1994). The interaction of phonetics and phonology in developmental verbal dyspraxia: Two case studies. *Clinics in Communication Disorders*, *4*(1), 66-77.

Velleman, S. L. (2002). Phonotactic therapy. *Seminars in Speech and Language, 23*(1), 43-56. doi: 10.1055/s-2002-23510

Wan, C. Y., Bazen, L., Baars, R., Libenson, A., Zipse, L., Zuk, J., ... Schlaug, G. (2011). Auditory-motor mapping training as an intervention to facilitate speech output in non-verbal children with autism: A proof of concept study. *PLoS ONE*, *6*(9), e25505.

https://doi.org/10.1371/journal.pone.0025505

Weismer, S. E., & Robertson, S. (2006). Focused stimulation approach to language intervention. In R. J. McCauley, & M. E. Fey (Eds.), *Treatment of language disorders in children* (pp. 175-202). Baltimore, MD, USA: Paul H. Brookes.

Woods, J. J., & Wetherby, A. M. (2003). Early identification of and intervention for infants and toddlers who are at risk for autism spectrum disorder. *Language, Speech, and Hearing Services in Schools*, *34*(3), 180-193. doi:10.1044/0161-1461(2003/015)

Yosick, R. N., Muskat, L. R., Bowen, C. N., Delfs, C. H., & Shillingsburg, A. M. (2016). Increasing single-word requests to multiword requests in children with autism and related disabilities. *Behavioral Interventions*, *31*(1), 28-43. doi: 10.1002/bin.1434

CHAPTER 3:

Skin Conductance as an *In-Situ* Marker for Emotional Arousal in Children with Neurodevelopmental Communication Impairments: Methodological Considerations and Clinical Implications [7]

Abstract

Even though electrodermal activity has been widely used in the study of psychological states and processes for over 130 years, the use of such technology *in situ,* within the context of daily activities, remains a major challenge. Recent technological advancements have led to the development of wearable biosensors that noninvasively measure electrical conductance across the skin. These biosensors represent a new approach for skin conductance assessment, as a proxy for emotional arousal, in children with neurodevelopmental communication impairments who are often described as having difficulties with emotional regulation, expressing thoughts and feelings, and present a higher prevalence of challenging behaviors. Here we provide an overview of skin conductance and explore the benefits of recent technological advancements for applied research and clinical practice. We draw on user experience from two experimental interventions involving eight children with neurodevelopmental disabilities. In both cases investigators monitored phasic and tonic EDA measures *in situ* using wearable biosensors. We share the behavioral and technical challenges experienced across these two experimental contexts, and propose associated considerations for future use. Specifically, sensor functioning, synchronization, and data preprocessing/analysis difficulties, as well as behavioral findings related to developmental differences, sensor tolerance over time, and sensor placement are discussed.

[7] The work in this chapter has been published: Aparicio Betancourt, M., DeThorne, L., Karahalios, K. & Kim, J. (2017). Skin conductance as an *in situ* marker for emotional arousal in children with neurodevelopmental communication impairments: Methodological considerations and clinical implications. *ACM Transactions on Accessible Computing*, *9*(3), 8. http://dx.doi.org/10.1145/3035536. LD & KK assisted with research design. LD edited prior drafts. JK analyzed EDA data for the OT-EDA study, and wrote the first draft of a subsection of the chapter included in section 3.2.5.2.2 Human computer interaction.

3.1 Introduction

Electrodermal activity (EDA) has been widely used in the study of psychological states and processes for over 130 years (Boucsein (1993) 2012) . However, traditionally its use has been limited to highly controlled settings in research laboratories with restricted external validity (e.g. Figner and Murphy 2011, Stevens and Gruzelier 1984, Williams et al. 2004). Recently, wearable biosensors that noninvasively measure electrical conductance across the skin have been introduced, offering the opportunity to study skin conductance *in situ*, within the context of daily activities (e.g. Poh et al. 2010a, 2012). Biosensors that measure EDA, among other biometrics, are already being commercialized and featured in mainstream media, with promises to "get the most out of your workouts" with the OM smart bra (OMsignal 2016), and to attain "a more calm, balanced state of mind" with Spire (Spire 2015). More relevant to this particular chapter, such technology has been marketed as able to predict "outbursts in individuals with autism" (Violeta 2015) (See also Curtis 2015, Kraft 2015, Stout 2015). The biosensor wristband is projected to "allow carers to monitor physiological signals that may be indicative of an impending meltdown, thereby allowing them ample time to take appropriate actions" (Violeta 2015). We contend that assessing skin conductance, as a proxy for emotional arousal, in children with neurodevelopmental communication impairments holds qualified promise in helping communication partners understand and interpret the perceptions and experiences of children with communication impairments. Despite promise, this technology is reaching the general public at a rate faster than associated guidelines for evidence-based practice.

The current chapter provides an overview of skin conductance and identifies the challenges associated with assessing skin conductance in children with neurodevelopmental communication impairments *in situ*, leading to the development of considerations for practice and the accumulation of resources to facilitate future data collection and analysis. The present work stems from our experiences assessing *in situ* skin conductance in children with neurodevelopmental communication impairments during two experimental contexts: 1) a speech-language intervention study focused on increasing multisyllabic productions in children with speech-language impairments (SL-EDA study) (see Aparicio Betancourt, DeThorne, and Karahalios 2013 for SL-EDA feasibility study, see DeThorne et al. 2015 for larger behavioral intervention project); and 2) an occupational therapy study focused on examining the use of a pressure vest to increase academic engagement for two children with intellectual disabilities (OT-EDA study) (Snodgrass et al. 2015).

We briefly review prior preliminary EDA findings from these two studies and also present novel post hoc data analyses specifically related to the utility of this technology with children with neurodevelopmental communication impairments in clinical settings.

3.2. Overview of Skin Conductance

EDA is a common term for the variation of electrical phenomena in the skin in response to sweat secretion, triggered by postganglionic sudomotor nerve fibers (i.e., sympathetic sweat motor nerve fibers, mediated by the neurotransmitter acetylcholine also referred to as cholinergic innervation) (Benedek and Kaernbach 2010a). The relationship between skin sympathetic nerve activity and electrodermal responses is complex (Kunimoto et al. 1992, Kunimoto et al. 1991), and is highly influenced by participant's thermoregulatory state (Wallin 1981). When comfortable ambient temperature is maintained (usually 22-24 °C), electrodermal activity has been shown to respond relatively slowly to sympathetic nerve impulses (i.e., usually 0.5-1.5 s and sometimes greater than 5 s), and the strength of the electrodermal responses also varies (Hagbarth et al. 1972, Kunimoto et al. 1992, Kunimoto et al. 1991, Wallin 1981). Differences in stimulus-response latency and sensitivity may be due to: interindividual differences in thermoregulatory state, differences in conduction time (e.g., conduction may vary based on the recording site and participant's height), the requirement of duct filling or sweat production potentiation before triggering an electrodermal response, and local stimuli including hormonal and mechanical stimuli (Kunimoto et al. 1992, Kunimoto et al. 1991, Wallin 1981). Although direct nerve recordings provide a more sensitive measure of cholinergic sympathetic activity (Brown et al. 2012), electrodermal activity remains a reliable and well-validated indirect measure which offers the opportunity to assess sympathetic nervous system (SNS) activity non-invasively and *in situ*. Broadly, EDA can be measured by placing two electrodes on the surface of the skin with or without applying a small external electrical current, referred to as exosomatic and endosomatic methods respectively. The use of EDA within psychological research has focused on exosomatic methods, recording skin resistance (R) or its reciprocal, skin conductance; R is equal to the voltage (V) applied between two electrodes placed on the surface of the skin, divided by the current (I) passed through the skin ($R = V/I$) (Boucsein (1993) 2012).

Changes in EDA may be caused by an increased or decreased demand for neural activity, including cognitive and emotional loads, and physical activity. Effortful allocation of attentional

resources, stress, affect, hydration of the corneum (upper epidermal layer), and diurnal effects (i.e., EDA levels are sensitive to the time of day), are factors known to influence EDA (Boucsein (1993) 2012, Critchley 2002, Dawson, Schell, and Filion (2000) 2007, Nagai et al. 2004). Specifically, an increase in cholinergic innervation (i.e., mediated by acetylcholine), as occurs for example in response to acute stressful stimuli such as being startled by a lightning bolt during sleep or in response to increased concentration due to increased task demands, leads to an increase in sudomotor SNS activity. This leads in turn to increased eccrine sweat gland secretion, and increased electrical or skin conductance that can be measured by biosensors (see Figure 3.1) (Benedek and Kaernbach 2010a, Squire et al. 2008). Baseline electrodermal activity may also vary with individual differences in race, age, sex, body mass index, sweat gland density, and with use of medications and psychoactive substances including caffeine (Doberenz et al. 2011). Electrodermal activity is measured as a one dimensional time series signal and consists of two main phenomena: phasic and tonic.

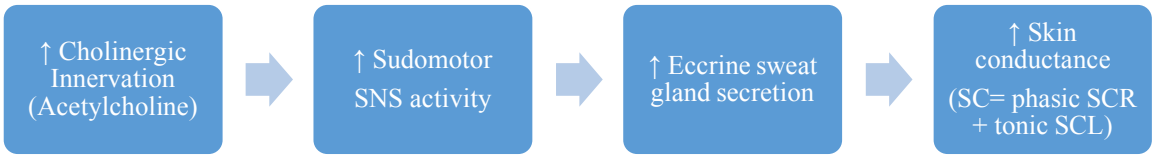


Figure 3.1. Sympathetic cholinergic (sudomotor) innervation predominantly mediates skin conductance.

3.2.1 Phasic skin conductance response. The phasic skin conductance response, reaction (SCR) or peaks, is a transient increase in skin conductance elicited one to five seconds after stimulus onset; novel, unexpected, significant, or aversive stimulus have been shown to produce an SCR (Boucsein (1993) 2012, Dawson, Schell, and Filion (2000) 2007). Consider, for example, an SCR elicited in a child who is first introduced to the shake of a tambourine, the sight of a toy snake, or the unexpected strobe light display at a museum. It is important to note, what may elicit an SCR in one person may not necessarily do so in another, and how our bodies respond to a stimulus may change over time or may vary based on context. For example, a self-identified autistic[8] woman, Dr. Temple Grandin, describes craving the "good feeling of being hugged", but

[8] We acknowledge the different perspectives surrounding identity-first and person-first language and use both as a means to recognize and appreciate the different opinions, and to improve sentence clarity (see Autism&Oughtisms 2011, Duncan 2011 for perspectives of parents of autistic children who use both identity and person-first language). Advocates of identity-first language propose autism is not a negative quality and is a part of the person, central to a person's identity, and should thus be used as an adjective (see Hillary 2015, Sinclair 1999 for perspectives written by autistic individuals). On the other hand, person-first language proponents view autism as a negative quality, as one of many traits of a person, that

then she describes how she "stiffened and pulled away to avoid the all-engulfing tidal wave of stimulation" (Grandin (1992) 2013). Temple Grandin may have experienced increased SCRs when hugged by others, compared to when she was at rest, and compared to a person who is not as sensitive to hugs.

In addition to specific phasic responses, there is spontaneous nonspecific phasic activity which cannot be linked to any specific stimulation. Nonspecific SCRs are most often used during continuous data collection, as it is difficult to link responses to specific stimuli. Nonspecific SCRs are considered a useful, indirect, index of sudomotor sympathetic activity, with higher frequency associated with higher levels of activity. The NS.SCR is frequently measured by both the amplitude and frequency of associated peaks, often ranging between 0.1-1.0 μS and 1-3 per minute respectively while at rest (Boucsein (1993) 2012, Dawson, Schell, and Filion (2000) 2007). In general, nonspecific SCRs can be observed while at rest due to spontaneous activity, and are also associated with relatively mundane stimuli within our day to day, as well as in response to less mundane stimuli such as the abrupt loud sound of a fire alarm. However, it may be difficult to link responses to specific external stimuli when collecting data in uncontrolled applied settings. Consequently, all SCRs gathered during continuous data collection, such as applied settings, are considered nonspecific.

3.2.2 Basal tonic skin conductance level. In contrast to the phasic component, the basal tonic skin conductance level (SCL) is relatively stable and associated with gradual changes in skin conductance. SCL will be relatively high in novel environments and will decrease gradually with time; it is associated with both cognitive and emotional arousal. As an example, the SCL of a child during their first day of school might be significantly higher compared to their SCL on subsequent days of the school year. The SCL has also been shown to be lower during sleep and higher during activated states such as states of increased concentration like when you are solving a math problem or learning to play the guitar. The SCL is frequently measured by the overall tonic level of electrical conductivity of the skin or as a gradual change measured at two or more points in time, often ranging between 2-20 μS and 1-3 μS respectively (Boucsein (1993) 2012, Dawson, Schell, and Filion (2000) 2007).

---

is not central to their identity, and believe using autism as an adjective leads to devaluation of the person and facilitates prejudice and discrimination (see Snow 2009 for a perspective of a mother of an adult with cerebral palsy).

3.2.3 Skin conductance measures. As aforementioned, phasic skin conductance response is frequently measured by both the amplitude and frequency of associated peaks. Basal, tonic skin conductance is frequently measured by overall skin conductance level, or change in SCL. Less common measures include: SCR area under the curve (AUC; i.e., the total area between the SCR initiation and recovery of SCR amplitude); and SCL AUC (i.e. the total area between a specified window of time) (Boucsein (1993) 2012, Dawson, Schell, and Filion (2000) 2007); see Figure 3.2 for an illustration of SC components and relevant measures). Correlations among EDA measures are generally moderate to low; however, the association between AUC and the other EDA measures is largely undocumented. Such findings suggest that most of the EDA measures represent partially independent sources of information, although the neurological and psychophysiological underpinnings are not well understood.

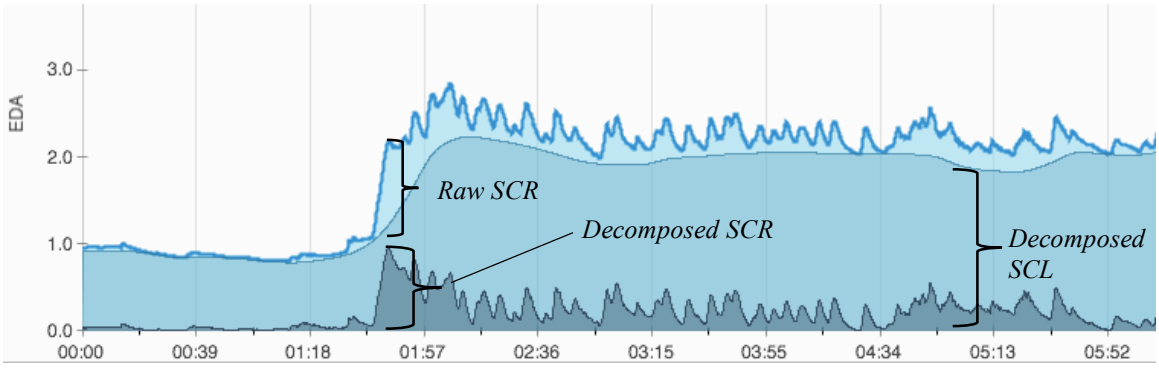


Figure 3.2. EDA/SC data as graphically displayed by BEDA (Kim et al. 2013). Light blue phasic SCR shows the SCR prior to SC decomposition (raw SCR). Dark blue phasic SCR and mid blue tonic SCL are the two components of SC post SC decomposition; SC= SCL + SCR.

3.2.4 Neural mechanisms of skin conductance. The neural mechanisms and pathways involved in mediating EDA are numerous and complex and further research elucidating the central origins of EDA is needed. Eccrine sweat glands have been shown to be mediated at different levels of the central nervous system, by both ipsilateral (i.e., same side of the body) and contralateral (i.e. opposite side of the body) processes, which are partly independent of each other. Boucsein ((1993) 2012) proposed three main central nervous system levels influencing EDA (see also Critchley 2002):

1) ***Ipsilateral Limbic Hypothalamic Source***: excitatory influences stem mainly from the hypothalamus, amygdala, and limbic system, whereas inhibitory influences stem from the

hippocampus. The limbic hypothalamic source is thought to be influenced by both thermoregulatory and emotional processes.

2) ***Contralateral Premotor Basal Ganglia Source***: independent of the first source and is composed primarily of excitatory and inhibitory influences originating on the premotor cortex, frontal cortex, and basal ganglia which mediate EDA in preparation of specific motor actions.
3) ***Reticular Formation (RF) Modulating System*:** both excitatory influences and inhibitory influences originate from the bulbar level of the RF on EDA, and may be eliciting or modulating influences that originate from the first two sources.

Generally, EDA during emotional tasks is thought to be mediated by the ipsilateral system, which may differentially influence eccrine sweat secretion (Mangina and Beuzeron-Mangina 1996), whereas EDA during nonemotional tasks is mediated by the contralateral system (Dawson, Schell, and Filion (2000) 2007). Differential influence of the ipsilateral system on eccrine sweat secretion has been supported by a direct electrical stimulation study in humans (Mangina and Beuzeron-Mangina 1996). SCRs elicited by direct stimulation of limbic structures, the amygdala in particular, via intracerebral electrodes in five adult neurosurgical patients, yielded a higher amplitude compared with those elicited from cortical areas, with higher ipsilateral than contralateral responses. Specific to the neural systems associated with the phasic skin conductance response see Critchley (2002), Dawson, Schell, and Filion ((2000) 2007), Mangina and Beuzeron-Mangina (1996), Nagai et al. (2004), and Vetrugno et al. (2003). Although there has been disproportionately less research examining the neurological underpinnings of the basal tonic skin conductance level readers are referred to Nagai et al. (2004).

3.2.5 Past and Common Uses. Recent advances in the development of biosensors have led to watch-like bands with embedded electrodes that noninvasively measure electrical conductance across the skin by passing a small amount of direct current between two electrodes in contact with the skin (bipolar recording) such as the Q sensors (Affectiva 2012) and E4 sensors (Empatica 2016b). In addition to measuring skin conductance (microsiemens, μs), such sensors may also measure temperature (Celsius), actigraphy (g-force, g), and photoplethysmography (which measures blood volume pulse from which cardiovascular measures such as heart rate can be extracted).

With the increased ease of EDA measurement has come a variety of new uses. In particular, EDA has been used in conjunction with functional magnetic resonance imaging (fMRI) to examine sympathetic activity during the decision-making process (Figner and Murphy 2011), as well as its association with implicit fearful experiences (Williams et al. 2004). One of its most promising areas is the potential to manage epilepsy (Empatica 2016a, Nagai and Critchley 2008) and predict seizures, particularly generalized tonic-clonic (GTC) seizures (Poh et al. 2012, Ramgopal et al. 2014).

In addition to such interesting venues, the use of EDA data may be particularly useful for better elucidating the everyday experiences of children with neurodevelopmental disabilities, especially those who experience significant communication difficulties. Although EDA varies widely across subjects, it is relatively stable within subjects, and changes in EDA within subject are associated with different psychological states (Boucsein (1993) 2012, Dawson, Schell, and Filion (2000) 2007). Consequently, monitoring EDA may assist in understanding children's response to various environmental conditions. One basic tenet, specified by the Yerkes Dodson Law (Yerkes and Dodson 1908), states that human performance increases as physiological arousal increases up to an optimal arousal point, beyond which it decreases; thus, representing a quadratic equation shown in an inverted-U curve. As such, EDA could help provide valuable information regarding how to best fashion educational environments to facilitate children's comfort and learning. More recently, researchers have also predicted it can be used to detect meltdowns in children with autism (Curtis 2015, Kraft 2015, Stout 2015, Violeta 2015). However, additional research is warranted investigating the use of biosensors to measure *in situ* skin conductance in applied settings.

3.2.5.1 Benefits of EDA assessment in children with neurodevelopmental communication impairments. Utilizing EDA to understand emotional arousal could be particularly useful for children who frequently encounter communication challenges. Neurodevelopmental communication impairments impact as many as 25% of children (Law et al. 2000) (See also ASHA 2014, Nelson et al. 2006, Pinborough-Zimmerman et al. 2007), including diagnoses of autism spectrum disorders, social (pragmatic) communication disorder, language delay or disorder, childhood apraxia of speech, and severe speech sound disorders.

Moreover, children with neurodevelopmental communication impairments are often described as having difficulties with emotional regulation and reported to have a higher prevalence of challenging behaviors. Redmond and Rice (1998) report a 50-70% co-occurrence between

language impairments and socio-emotional difficulties, most likely due in part to communication challenges. Challenging behaviors include both externalizing behavioral problems such as physical aggression (e.g., hitting, kicking, biting), verbal aggression, oppositional behaviors (e.g. running away), and internalizing behavioral problems such as anxiety, depression, and social withdrawal (cf. Qi and Kaiser 2004). Studies suggest that interventions aimed at understanding the feelings and intentions of individuals with communication challenges are likely to decrease the frequency and severity of challenging behaviors (Gainey 2013, Halle, Ostrosky, and Hemmeter 2006, Hemmeter, Ostrosky, and Fox 2006, Hutchins and Prelock 2014). Our hope is that measures of physiological arousal could aid in better understanding the emotional arousal levels and corresponding emotions of children with neurodevelopmental communication impairments in order to build better environments to support their learning and social interaction.

3.2.5.1.1 Sensory integration differences in children with autism. Electrodermal activity assessment might be particularly useful for children with communication and sensory integration difficulties, such as many children with autism. In fact, hyper/hypo-reactivity to sensory input or unusual interest in sensory aspects of the environment is a diagnostic feature for autism spectrum disorders as detailed in the Diagnostic and Statistical Manual of Mental Disorders, fifth edition (APA 2013), and is commonly reported by autistic adults (Grace 2015, Fleischmann and Fleischmann 2012). Stevens and Gruzelier (1984) assessed EDA (i.e., SCL and SCRs) in a controlled laboratory setting and demonstrated that levels of skin conductance were slightly higher in autistic children compared to neurotypical children and those with intellectual disabilities (ages 7-17 years) in response to auditory stimuli, thereby providing evidence for a tendency toward heightened physiological arousal. In addition, the children with autism showed reductions in peak amplitude, longer response latencies and rise times in response to 70 dB tones, indicating a slower habituation and relative delay in stimulus registration (See also Ming et al. 2005). Moreover, differences in skin conductance in response to another person's gaze have also been reported in children with autism, with increased SCRs associated with straight-forward gaze compared with averted gaze (Kylliainen and Hietanen 2006). Differential pattern of SCRs to the two gaze conditions was not seen in children without autism. Together, such results support the sensory integration differences often reported in individuals with autism (Tomchek and Dunn 2007), and suggest that EDA data might be useful in adjusting intervention environments.

3.2.5.2 In situ EDA assessment.

3.2.5.2.1 Behavioral sciences. Despite the recent technological advances in EDA, few studies have examined *in situ* EDA in children with communication impairments within the behavioral sciences, and the impact of intervention on children´s emotional arousal is relatively unexplored. Hedman (2010) was the only study we found that examined EDA during intervention, and it was largely observational in nature. Specifically, EDA was measured using a biosensor known as iCalm in 22 children with sensory processing disorder during particular guided occupational therapy activities. Each child participated in three to eight guided activities per session and analyses were based on approximately 50 total one-hour sessions. The lack of an experimental control, the lack of normalization of the data and the low power led to inconclusive results. However, the researchers highlighted results from individual children in response to specific stimuli. For example, one child demonstrated increased emotional arousal when playing in the ball pit, whereas another child demonstrated significant decreases during the same guided activity. The study highlighted the importance of case study methodologies that can assess and accommodate children's individual differences.

Similar to Hedman (2010), Miller, Coll, and Schoen (2007) examined EDA in children undergoing occupational therapy. However, skin conductance was not assessed during the intervention; instead, SC was continuously recorded during the Sensory Challenge Protocol before and after treatment. The Sensory Challenge Protocol consisted of a series of 50 sensory stimuli administered to the participants. Miller, Coll, and Schoen (2007) conducted a pilot randomized controlled trial in 24 children with sensory modulation disorders to assess the effectiveness of occupational therapy using a sensory integration approach. Even though group comparisons had limited power, were nonsignificant, and 54% of the EDA data had to be discarded, their results indicated greater reduction in SCR amplitude in children in the treatment condition compared with the children in the active and non-active control groups suggesting reduced hyper-reactivity.

3.2.5.2.2 Human computer interaction. Although a large number of studies have investigated ways to automatically identify individuals' stress levels (e.g. Healey and Picard 2005, Sano and Picard 2013), engagement levels (e.g. Hernandez et al. 2014), and various emotions such as happiness (e.g. Jaques et al. 2015) using EDA and other physiological data such as heart rate in individuals without communication impairments, fewer studies within computer science have explored the use of such systems in individuals with neurodevelopmental communication

impairments. More recently, several technological applications have been developed to measure EDA in conjunction with eye gaze measures and measures of social communication in autistic individuals, to better understand how their physiological levels change when demonstrating challenging behaviors or during social interactions (el Kaliouby and Goodwin 2008, Lee et al. 2008, Riobo et al. 2014). The goal of these systems is to support communication by providing caregivers of individuals with autism and autistic individuals themselves with alternate methods to observe their levels of emotional arousal in real-time during social interactions. Reviewing EDA levels synchronized with video data of challenging behaviors or day to day social interactions could potentially be used to provide biofeedback to autistic individuals to facilitate processing social information and facilitate regulating emotional arousal levels prior to engaging in challenging behaviors.

For example, Lee et al. (2008) developed a wearable platform to detect face contact using a hat-mounted wireless camera while measuring EDA via a wrist-worn sensor. The researchers conducted a usability study with four autistic adolescents and their caregivers to explore how the system could be improved to be used to quantify social stimuli and the associated stress response, as measured by skin conductance. Findings illustrated various limitations of the system including the need to develop a more flexible system that provides information in real-time. In fact, one of the caregivers who participated in the study suggested that real-time visualization of physiological states of people with autism "would be helpful in understanding students' arousal states and in teaching self-regulation" (Lee et al. 2008). With a similar motivation, the Interactive Social-Emotional Toolkit (iSET) (el Kaliouby and Goodwin 2008, Madsen 2010) was developed by combining a wearable camera, that can be worn as a self-cam or a head-cam, and a wrist-worn sensor to capture video, audio, and physiological data (i.e. skin conductance, heart rate, and movement) concurrently. Young autistic adults have participated in the early design stages and used prototypes of the wearable system that captures multimodal data. The goal of this technology is to facilitate the processing of high-speed and complex social information such as nonverbal cues, by providing autistic adults with the opportunity to systemize, quantify and reflect on their own social interactions via a fun and engaging tool. Finally, Riobo et al. (2014) developed a system that sent a real-time visualization of an autistic child's EDA to the Google glass unit (*Google Inc*.) worn by their caregiver, while the two were interacting together. The system aims to help caregivers interact with their children by gaining a deeper understanding of the child's internal

state and their individual experience of the social interaction. These different technologies however, have yet to be evaluated in a large-scale basis or in in-depth qualitative, descriptive studies evaluating user experience.

Although technological applications and computational methods have been developed and have shown promise of using EDA to understand and predict behaviors in individuals whom are often difficult for their partners to understand, there is limited research investigating and/or reporting the challenges in using wearable sensors to assess *in situ* EDA in children with neurodevelopmental communication impairments (cf. Boucsein (1993) 2012, Doberenz et al. 2011, Turpin, Shine, and Lader 1983). For example, wearing a novel object, such as a sensor, could be a challenging task in itself for children with autism and/or sensory difficulties. In this way, we aim to provide methodological considerations for assessing short-term *in situ* skin conductance based on investigator experience working with children, ages 2-11, with neurodevelopmental communication impairments across two applied experimental contexts.

3.3 Experimental Context

This section introduces the two experimental contexts used to inform our considerations for use of biosensors *in situ* with children who have neurodevelopmental communication impairments: skin conductance assessment during speech-language therapy, followed by a study assessing skin conductance during occupational therapy. Specifically, we provide a brief overview of the purpose, method, and results associated with each study, with an explicit focus on novel post hoc data analysis related to use of the biosensors to collect EDA during behavioral intervention.

3.3.1 SC assessment during speech-language therapy. This feasibility study, hereafter referred to as the SL-EDA study, assessed the associations between *in situ* EDA recordings and off-line behavioral coding of emotional valence (EV) and examined the association among different EDA measures (Aparicio Betancourt, DeThorne, and Karahalios 2013)[9]. The study represents a subset of participants (n= 6) from a larger project that examined the effectiveness of a speech-language intervention for children at the single-word stage of development (i.e., DeThorne et al. 2015). Participant demographic and novel behavioral data regarding biosensor use within the SL-EDA study is presented in Table 3.1 by participant (identified by pseudonyms). The

[9] SCL. AUC is referred to as SC.AUC in Aparicio Betancourt et al., 2013 but should read as SCL.AUC

first author served as one of the intervention therapists for all participants except Karis, and all sessions were video- recorded. Note that Dora and Pan withdrew early from the larger intervention study, and Heidi never tolerated wearing the biosensors[10].

Table 3.1. *SL-EDA Participant Demographic & Biosensors Information*

Participant Pseudonym	Gender	Age	Race	Primary Clinical Diagnosis	Q Sensor Placement	Total # of sessions [a]	# Desensitization Sessions [b]	% Sessions Sensor Worn [c]
Pyrros	Boy	2;7	White	Unspecified Developmental Delay	Left & Right Ankle	37	1	97%
Dora	Girl	2;11	Asian	Autism	Left & Right Wrist/Ankle	9	0	89%
Angelo	Boy	3;0	White	Speech Sound Disorder	Left & Right Ankle	24	8	67%
Heidi	Girl	3;8	White	Autism	NA	25	25	0%
Karis	Boy	4;10[d]	White	Autism	Right Wrist	20	5	60%
Pan	Boy	4;11	White	Autism	Left & Right Ankle	3	2	33%

[a] Total # of sessions refers to the total number of sessions, including both treatment (tx) and evaluation sessions, in which the participant was exposed to the sensor(s).[b] # Desensitization Sessions refers to the number of sessions the child was exposed to the sensor(s) prior to wearing it for at least half the tx session (i.e., 22.5 minutes). [c] % Sessions Sensor Worn refers to the percentage of sessions at least one sensor was worn for at least half the tx session (i.e., 22.5 minutes). [d] Karis' age was reported as 4;8 in Aparicio Betancourt et al., 2017 but should read as 4;10 (i.e., 4.8 years).

EDA was measured in microSiemens (μS) and recorded from dry Ag/AgCl disk electrodes at a sampling rate of 8 Hz using the Affectiva Q sensor v1 as long as children were willing to wear the biosensors (see Figure 3.3, left panel). The biosensors were fitted at the wrist or ankle based on participant tolerance and preference was given to the dominant wrist/ankle as it yielded a higher amplitude compared to the nondominant wrist/ankle (see Román et al. 1989 for similar results, see Picard, Fedor, and Ayzenberg 2016 for significant EDA asymmetry, with greater amplitude in dominant wrist); for consistency, only data from the dominant wrist/ankle were analyzed. Although EDA data collected at the wrist yielded slightly higher responses compared to the ankle, sensor location attempted to increase participant's comfort level and decrease distractions (see Payne et al. 2013, for similar results). Collection of EDA data at both the wrist and the ankle do

[10] Dora's family moved out of state and Pan's mother reported that he was already overcommitted with other activities.

not interfere with daily activities and have been shown to be accurate and strongly correlated with more traditional palmar sites of EDA measurement (Fletcher et al. 2010, Picard, Fedor, and Ayzenberg 2016, Poh, Swenson, and Picard 2010b). The biosensors were calibrated to be the same within study. The number of sessions participants were exposed to the sensor(s) prior to wearing them for at least half the treatment session (i.e., # desensitization sessions), and the length of the desensitization procedures varied by participant based on sensor tolerance (see section 3.5.1 for a review of the desensitization techniques). Even though dry electrodes were used, to improve signal acquisition electrode gel was placed on each electrode, and children participated in a biosensor acclimation process which involved a period of activation followed by a rest period intended to bring the children's physiological arousal level back to a theoretical baseline. Following the biosensor acclimation process, children participated in a motor practice session that involved repetitions of multisyllabic speech productions (e.g. *butterfly*, *tiger*), and a developmental play session in which those same words were modeled and elicited during child-centered play-based activities (Figure 3.3) (See DeThorne et al. 2015 for additional treatment details). The average session length for the SL-EDA study was 55 minutes (i.e., 10 minutes for the biosensor acclimation process and 45 minutes for therapy).



Figure 3.3. Left panel: Affectiva Q Sensor (Affectiva 2012). Mid panel: Child practicing target words with VocSyl, a software developed to provide online visual feedback during the motor practice portion of the intervention; biosensors are placed on his ankles and are not visible. Right panel: Child engaged in a naturalistic interaction during developmental play targeting the disyllabic word, *tiger*; biosensors are placed on his left and right ankles.

Consistent with prior literature (Boucsein (1993) 2012, Dawson, Schell, and Filion (2000) 2007, Hernandez et al. 2014), the EDA measures assessed included nonspecific skin conductance response frequency (NS.SCR.freq) per minute, nonspecific skin conductance response amplitude

(NS.SCR.amp), and an additional less traditional measure, nonspecific skin conductance level area under the curve (SCL.AUC). SCRs were detected using an amplitude threshold ≥ .01 μs and a minimum distance between responses of at least 1 second. SCL.AUC was calculated using one-minute consecutive windows from the start of the session to the end of the session. The thermostats at the clinical facility were set at a comfortable ambient temperature, approximately 71 °F (22 °C), and all analyzed intervention sessions were held at the same time each day within participant. EDA data were analyzed for participants who completed the speech-language intervention and wore the biosensors (i.e., Pyrros, Angelo, Karis). Although EDA data were collected across both treatment and evaluation sessions, data were analyzed for treatment sessions only due to differences in activity. Data from 33% (21/63) of all treatment sessions had to be discarded due to a) sensors not worn or tolerated for at least half the session (11/21, 52%), b) low signal to noise ratio associated with loss of electrode contact with the skin due to physical activity, such as jumping and hand flapping, and participant manipulation with the sensors (5/21, 24%), c) malfunctioning sensors or synchronization difficulties (4/21, 19%), and d) sessions being held at a different time of the day (1/21, 5%). The percent of sessions with useable EDA data by child was 91% (29/32) for Pyrros, 61% (11/18) for Angelo, and 15% (2/13) for Karis. Consistent with signal processing and continuous skin conductance processing (Benedek and Kaernbach 2010b, Boucsein (1993) 2012, Dawson, Schell, and Filion (2000) 2007, Hernandez et al. 2014), EDA analyses followed a 3-step process: 1) visual inspection of 10% of synchronized treatment sessions with EDA recordings using ELAN[11]. 2) data pre-processing (visual inspection of all EDA recordings, cropping, manual artifact rejection and smoothing), and 3) SC decomposition into continuous phasic SCR and tonic SCL components using BEDA [12] (Kim et al. 2013) and MATLAB. Actigrahy (3-axis accelerometry) and skin surface temperature data were used to aid synchronization, visual inspection of the data, and manual artifact rejection. In addition to assessing emotional arousal, emotional valence was assessed by conducting off-line behavioral coding of all treatment sessions for which EDA data were analyzed. Specifically, emotional valence was rated every minute based on an examiner's video review of the child's vocalizations, facial expressions, and corporal gestures, using a 1-5 Likert scale (1 = high negative affect, 5 = high positive affect).

[11] ELAN is an annotation tool that allows you to converge multi-media recordings (Brugman and Russel 2004, MPI 2016, see Berez 2007 for a review)

[12] BEDA is a visual analytic tool to help synchronize, visualize, and analyze data sources for multiple sessions of behavioral and physiological data.

As a brief overview of results from the SL-EDA study, associations between *in situ* EDA and off-line behavioral coding of emotional valence were examined by polynomial regression for all treatment sessions. NS.SCR.freq was significantly higher for high positive and negative off-line behavioral coding of emotional valence ratings compared to neutral ratings ($b = .53$, $p < .01$) though the effect size was small (adjusted $R^2 = .023$, $F(4, 1{,}337) = 8.78$, $p < .001$). Although significant, NS.SCR.freq explained very little variance in behavioral ratings of emotional valence (< 3%). Consistent with the general trends recorded in other studies for the average physiological responses for children who were more difficult to engage (Hernandez et al. 2014) or for children introduced to novel stimuli (Dawson, Schell, and Filion (2000) 2007), both NS.SCR.freq ($b = .14$, $p < .001$) and SCL.AUC ($b = .20$, $p < .001$) significantly increased over time within session though with fairly small effect sizes[13] (NS.SCR.freq: adjusted $R^2 = .019$, $F(2, 1{,}339) = 13.66$, $p < .001$; SCL.AUC: adjusted $R^2 = .053$, $F(2, 1{,}339) = 38.75$, $p < .001$). In regard to associations across EDA measures, consistent with prior research, results indicated a significant moderate correlation between NS.SCR.freq and NS.SCR.amp ($r = .48$, $n = 42$, $p < .01$). In contrast, the association between SCL.AUC and other EDA measures is largely undocumented. Results indicated a significant high correlation between NS.SCR.freq and SCL.AUC ($r = .85$, $n = 42$, $p < .001$), and a significant moderate correlation between NS.SCR.amp and SCL.AUC ($r = .44$; $n = 42$, $p < .01$), suggesting SCL.AUC and NS.SCR.freq may be mediated by similar psychophysiological or neurological sources.

3.3.2 SC assessment during occupational therapy. The second experimental context reviewed here is a single-subject reversal design study (A-B-C-A) focused on examining the effects of a pressure vest on academic engagement, challenging behaviors, and skin conductance in two children with neurodevelopmental disabilities (Snodgrass et al. 2015), hereafter referred to as the OT-EDA study. The use of pressure vests and other sensory integration techniques are common practice within occupational therapy for individuals with increased needs for proprioceptive/tactile input as a means to help regulate physiological arousal (Barton et al. 2015, Hodgetts, Magill-Evans, and Misiaszek 2011, Lang et al. 2012, Reichow et al. 2009). The use of biosensors was included to assess EDA as a dependent variable, in addition to behavioral

[13] Although the general protocol of the speech-language treatment sessions remained the same, children engaged in a variety of activities during the latter half of the session consisting of developmental play and were therefore not habituated to the stimuli (see DeThorne et al. 2015 for treatment details).

measures. Participant demographic and novel behavioral data regarding biosensor use within the OT-EDA study is presented in Table 3.2 by participant (identified by pseudonyms). All sessions were video-recorded.

Table 3.2. *OT-EDA Participant Demographic & Biosensors Information*

Participant Pseudonym	Gender	Age	Race	Primary Clinical Diagnosis	Q Sensor Placement	Total # of sessions [a]	# Desensitization Sessions [b]	% Sessions Sensor Worn [c]
Damien	Boy	9;8	Black	Intellectual Disability (ID) [d]	Left & Right Ankle	37	0	100%
Calvin	Boy	11;6	White	Autism, ID	Left Ankle	23	0	100%

[a] Total # of sessions refers to the total number of sessions, including both observation and treatment (tx) sessions, in which the participant was exposed to the sensor(s). [b] # Desensitization Sessions refers to the number of sessions the child was exposed to the sensor(s) prior to wearing it for at least half the tx session (i.e., 7 minutes). [c] % Sessions Sensor Worn refers to the percentage of sessions at least one sensor was worn for at least half the tx session, (i.e., 7 minutes). [d] Damien had secondary diagnoses that included physical impairment, complex partial seizures and absence seizures.

In contrast with the SL-EDA study, the OT-EDA study collected electrodermal activity at a sampling rate of 32 Hz using the Affectiva Q sensor v2. In addition to the functions present in Affectiva Q sensor v1 used by the SL-EDA study, v2 offered wireless capabilities via Bluetooth connection. The sensor could transmit data between the Q sensor and a computer in real-time for visualization using Q Live Software. However, only data from the internal card data were analyzed (see section 3.4.2.1 for further details). The biosensor(s) were fitted on the ankle and secured with an ankle wrap. Similar to the SL-EDA study, preference was given to the dominant ankle and only data from the dominant ankle were analyzed. Sessions consisted of a biosensor acclimation process followed by occupational therapy activities individualized to each participant's goals, such as fine motor skills (e.g., self-feeding with a fork) and cognitive skills (e.g., counting, sorting objects by attributes such as color) (Figure 3.4) (see Snodgrass et al. 2015 for additional treatment details). The reversal-design (A-B-C-A) consisted of three main conditions: unpressurized vest, structured teaching (condition A); unpressurized vest, unstructured teaching (condition B); and pressurized vest, unstructured teaching (condition C). The average session length for the OT-EDA study was 24 minutes (i.e., 10 minutes for the biosensor acclimation process and 14 minutes for therapy).



Figure 3.4. Left panel: Child engaged in a communicative exchange, using his speech-generating device, prior to the treatment session; biosensor is placed on his left ankle and secured with an ankle wrap. Right panel: Child putting toys into a container targeting fine-motor skills during occupational therapy; biosensors are placed on his ankles and are not visible.

Consistent with prior literature (Boucsein (1993) 2012, Dawson, Schell, and Filion (2000) 2007, Hernandez et al. 2014), the EDA measures assessed included mean NS.SCR.freq per minute per session, mean NS.SCR.amp per session, and mean skin conductance level (SCL) per session. SCRs were detected using an amplitude threshold ≥ 0.05 µs and a minimum distance between responses of at least 3 seconds. Mean SCL measures an average of all the collected SCL values (32 hz per second) across a treatment session. Mean SCL was highly correlated with mean SCL.AUC (r = .90). All analyzed intervention sessions were held at the same time each day within participant. EDA data were analyzed for the one participant who reached a stable baseline for behavioral measures and wore the biosensors (i.e., Damien). Although EDA data were collected across both observation and treatment sessions, data were analyzed for Damien's treatment sessions only, for a total of 25/30 (83%) treatment sessions, as observation sessions were conducted for desensitization purposes. The observation sessions were used to gradually pair the sensors with feelings of relaxation or excitation prior to the onset of treatment sessions to accustom children to wearing the sensors. Data from the remainder 17% (5/30) of the sessions had to be discarded due to a) sessions ending early due to health concerns related to Damien's seizure diagnoses (3/5=10%) [14], and b) low signal to noise ratio associated with extremely low electrodermal responses (< 0.7 µS) based on a criterion of one standard deviation away from the

[14]Damien also presented seizure activity before and/or during 10 out of the 25 treatment sessions analyzed. These sessions were not discarded because Damien was able to safely participate for the majority of the duration in each of these treatment sessions. Although SCL (M = 4.16 µs, SD = 1.93 µs), NS.SCR.amp (M = 0.09 µs, SD = 0.07 µs) and NS.SCR.freq per minute (M = 0.23, SD = 0.2), were lower for the seizure days than for the non-seizure days (SCL: M = 4.28 µs, SD = 1.92 µs; NS.SCR.amp: M = 0.15 µs, SD = 0.15 µs; NS.SCR.freq per minute: M = 0.37, SD = 0.20), a *Mann-Whitney U test* indicated that scores were not significantly different.

mean (2/5 =7%). EDA data analysis followed a similar 3-step process compared with the SL-EDA study with one major difference: all of the video-recorded treatment sessions were synchronized with EDA recordings and visually inspected using BEDA compared to the 10% of the sessions in the SL-EDA study (see section 3.3.1).

As a brief overview of results from the OT-EDA study, the interventionist's instructional practices (i.e., structured v. unstructured) appeared to play a more direct role in child engagement for Damien than did use of the pressure vest based on visual inspection as is consistent with single-case design. Specifically, Damien remained engaged approximately 85% of the time and presented challenging behaviors approximately 10% of the time during the structured instruction phases, compared to 34% and 69% respectively during the unstructured instruction phases. In addition, use of the pressurized vest was not significantly associated with EDA based on *Mann-Whitney U tests* across conditions, and associations between EDA levels and child engagement or challenging behaviors were inconclusive. Readers are referred to (Snodgrass et al. 2015) for a more detailed account of the findings.

3.4. Methodological Considerations

3.4.1 Behavioral findings across studies. The present section summarizes children's responses to use of the biosensors across both the SL-EDA and OT-EDA studies. Given the age and linguistic ability of the participants, we relied on video-recorded behavioral data from the participants and direct interviews from the research assistants regarding their experience with the technology. In particular, an additional coding pass was conducted across available video data in order to review participant's behavioral response to wearing the Q sensor(s). The first author conducted an explicit review of at least 10% (n = 3) of all sessions for each child across the two studies (Range = 11% - 100%), including the first two sessions and the last session. Video-recorded sessions included observations, evaluations and treatment sessions. Additionally, to gain better insight on the user experience of the investigators, the first author conducted semi-structured interviews with the three lead research assistants across the two studies face-to-face or by telephone. Data were analyzed by the first author following qualitative thematic methodology guidelines (Braun and Clarke 2006) including familiarization, development of themes, and developing an analytic narrative. Observation notes were taken during the interviews and all interviews were audio-recorded and transcribed. Orthographic transcription consisted of three

levels of coding: the first two levels were conducted by an individual transcriber trained to replicate pauses, non-speech sounds, overlapping speech, and unintelligible speech; the second level allowed the transcriber to verify the accuracy of the original transcription. The purpose of the third coding level was to reach consensus across two transcribers on the transcript sections with unintelligible speech. After becoming familiar with the interview data, the first author conducted a preliminary analysis of the data to identify codes and themes among the data related to behavioral challenges, technical challenges, and recommendations. Finally, the first author reviewed the preliminary analysis and re-reviewed all interview transcripts to identify themes and extract representative excerpts across the aforementioned categories.

Despite substantial individual variability, three noteworthy themes emerged in regard to behavioral challenges associated with sensor use: a) developmental differences, b) sensor tolerance, and c) sensor placement.

3.4.1.1 Developmental differences. First, younger children required a period of acclimation/desensitization to the biosensors before wearing them. All the children below the age of 5, with the exception of Dora, required at least one desensitization session (Median = 4) whereas our two oldest participants across studies, ages 9 and 11, wore the sensors during the initial session and did not show any visible signs of discomfort. Dora was our first participant across both the SL- and OT-EDA studies and was a child who was largely nonverbal. She appeared highly reluctant to wearing the sensors at first. When initially introduced to them, she pulled her hand away and squirmed to avoid having the sensor placed on her wrist. “Just a little bit. Ok?” her mom remarked. Once the sensor was on her wrist, as quickly as possible, the clinicians diverted her attention by inviting her to bounce on a green stability ball. As Dora’s facial expression began to relax and once again smile, the clinicians resumed placing the other sensor on. Once again, Dora began to squirm and then vocalize, and finally tried removing the sensors with her mouth. However, as soon as her attention was diverted away from the sensors and towards the stability ball, Dora instantly engaged in the ball activity by constant laughing as the clinicians and her mother sang, “bounce, bounce Dora…Dora likes to bounce!”

3.4.1.2 Sensor tolerance. A second notable finding is that despite initial reluctance by many of our younger participants, ages 2 to 4, the majority of children started wearing the sensors consistently in subsequent sessions (Median = 61% of sessions). Our two older participants, ages 9 and 11, wore the sensors consistently across all sessions (Median = 100%). For example, from

the second session onwards Pyrros began to spontaneously ask for the sensors "papa watch?" and to voluntarily participate in the process of putting the sensors on and placing them under his socks. Angelo's reaction to the sensors went from abrupt, unanticipated crying in the initial session to wearing the sensors for the entire 45 minutes of the ninth treatment session. As described by Angelo's lead clinician when asked to describe the process of using the sensors in the SL-EDA study, "I know for a second [Angelo] was getting freaked out and started screaming and crying and wouldn't have anything to do with [the sensors], but after that…one of the first things we did was…let him work out which one he wanted to put on first….and then we would strap 'em on there…and then he would wear them for the duration of the session".

Despite such success, there were two notable exceptions to the developed tolerance of the sensors over time: Karis and Heidi. Karis, as described by his mother was "obsessed" with watches, which hold close resemblance to the Q sensor. Karis enjoyed exploring the sensor and placing it on and taking it off his wrist starting with the second session. He showed a preference for placing the sensor on his wrist, and did not tolerate the sensor on his ankle. Unfortunately, Karis' interest in the sensor distracted him during therapy, and it was often put aside by the clinicians to be able to target the speech and language treatment goals. As reported by Karis' clinician when prompted to describe the children's response to the sensors, she explained, he "was just distracted by [the sensor], he kinda just wanted to play with it…he liked to put the watch on himself…and he always wore [it] pretty loose too…that sometimes was a challenge when he was struggling with it and you know we needed to get on…with the treatment session". Because of the constant manipulation of the sensor, paired with repetitive motor mannerisms (e.g., arms and hand flapping and posturing, and up and down bouncing), a high percentage of the EDA data collected during the times he wore the sensor for at least half of the tx session had to be discarded due to low signal-to-noise ratio (SNR) (5/8 = 63%). That is, he wore the sensor for 22.5 minutes in 8 of his 13 tx sessions, and 5 of those 8 tx sessions were discarded due to low SNR, even though he wore at least one sensor for at least half the session for 12 of his 20 sessions (including both tx and evaluation sessions), and for all of his sessions in the second half of the intervention (i.e., session 11-20). Heidi on the other hand, did not seem to trust the sensors. Even when the sensors were placed on family members she seemed protective of her loved ones and would immediately remove the sensors from them. Through a gradual desensitization process, she allowed the sensors to be placed on her ankles for a few seconds but would quickly remove them. Although a variety of strategies were used as an

attempt to have Heidi wear the sensors during therapy, including a social story (see Appendix A), Heidi never did tolerate wearing the sensors on either the wrist or ankle for an extended period of time; she did not wear either sensor for at least half the session (i.e., 22.5 minutes) for any of her 25 sessions.

3.4.1.3 Sensor placement. A third and final behavioral trend worth noting is that the sensors were better tolerated on the ankle secured with either an ankle wrap or socks than on the wrist. In addition to changing the 'feel' of the biosensor, such cover might also make it more likely that children forget they have it on, decrease the likelihood of children becoming distracted by the sensor throughout the session, as well as decrease movement artifacts associated with loss of electrode contact with the skin, often caused by physical activity (such as walking, running, or other repetitive motor movements). In the words of one of the SL-EDA clinicians, once the sensors were on the children's ankles, "the kids could put them under their socks…you know they kinda would forget about them and not be messing with them or be distracted by them." In Dora's case, placing the sensors out of sight, under Dora's sleeves and particularly under her socks and pants was helpful; Dora seemed to barely notice them in such cases. She went from squirming to completely disregarding the sensors when placed on her ankles instead of on her wrists. Given that Dora tolerated the sensors on her ankles more so than on her wrist; we continued to attempt to place the sensors on her ankles and opted to do the same for all subsequent participants. We also encouraged caregivers to dress their child with long pants or sleeves. Ultimately, sensor location varied by participant based on individual participant's preference and tolerance.

Even though the investigators moved towards placing the sensors out of sight, many of the younger children, including Angelo and Karis, often became distracted by the intermittently flashing light displayed on the sensor. This led them to manipulating the sensor, increasing the artifacts in the data, and on occasion the children removed the sensors. Instead of relying solely on whether or not the participants wore socks and/or long pants during the session to hide the sensors from sight, whenever the child was not wearing socks or to avoid having to tighten the sensors too much due the short length of the sensor strap, the clinicians in the OT-EDA study secured the sensors with an ankle wrap.

In sum, children tolerated the sensors better when placed on the ankle. Although younger children required a period of desensitization to the biosensors before wearing them, most of our

participants started wearing the sensors consistently in later sessions and older children consistently wore the sensors throughout all of the sessions.

3.4.2 Technical findings across studies. Similar to the behavioral findings, noteworthy themes related to technical challenges were derived from behavioral observation of the video-recorded sessions and the semi-structured interview data. In support of the point that technical challenges were significant and impinged on the data collection, 4% (4/93 sessions) of the data were discarded due to technical challenges, keeping in mind there would have been substantially more data loss if only one sensor would have been available during data collection. Given two sensors were available during data collection, researchers were able to collect data when one of the two sensors malfunctioned (i.e., if the dominant wrist/ankle sensor malfunctioned, the second sensor could be placed in the dominant wrist/ankle in order to collect data). Moreover, all investigators, including those with prior signal processing experience, reported encountering several technical challenges during the research process.

Across the two experimental contexts, the following three main themes associated with technical challenges emerged: a) sensor functioning, b) synchronization, and c) data preprocessing/analysis. Of interest, 3 sessions were discarded due to sensor functioning and 1 session was discarded due to synchronization challenges.

3.4.2.1 Sensor functioning. Specific to sensor functioning, investigators had difficulty getting the sensors to begin recording data after fitting the sensors on the child during the initial evaluation sessions. Based on the Affectiva Q User Manual (2013) the "Q Sensor is ideal for long term use because it works without the use of gels"; instead, Q sensors use dry electrodes. However, without the use of electrolyte, even after engaging in a task that stimulated cognitive, emotional and physical activation, the sensors would not start logging data automatically or manually for the first few minutes. As a result, in order to improve signal acquisition investigators in the SL-EDA study continued using an electrode gel as the conductive medium between the electrodes and the skin during all treatment sessions. One of the lead clinicians in the SL-EDA study explained, "to get the sensors to conduct" we "figure[d] out to put just a little bit of gel on there…" Investigators in the OT-EDA study subsequently adopted this practice as well. Moreover, OT-EDA investigators also experienced difficulties visualizing and recording data in real-time via Bluetooth. The real-time data were not reliable because a) the Bluetooth was often disconnected, and b) data resulted in a drift.

In addition to difficulty with data recording, technical challenges related to sensor functioning also included accessing the data and sensor malfunctioning. Specifically, investigators across both experimental contexts had difficulties with Windows computers failing to recognize the sensor as an external device and as a result were initially unable to access the data. Upon contacting customer service, investigators in the SL-EDA study were informed this occurred due to "the number of directories in the root exceed[ing] a threshold" of approximately "20 dated folders" (Affectiva Support, pers. comm.). Also across the course of both studies, one of the two sensors stopped recording data and required repair. Although the reason for malfunctioning in the OT-EDA study was never resolved, the device in the SL-EDA was shipped back to the manufacturer for repair in the midst of data collection. The reason for the malfunctioning noted by the manufacturer was that the sensor's secure digital (SD) memory card became dislodged from its housing (Affectiva Support, pers. comm.).

3.4.2.2 Synchronization. The second theme of technical challenges associated with use of the sensors across both studies related to synchronization of electrodermal activity with audio-video recordings. As specified by one investigator from the OT-EDA study when asked about the technical difficulties encountered during the study, "the **huge** difficulty was to synchronize the video and the sensor data streams because the [camera] and the sensor technology work separately so we had to time it and [use] different software programs to synchronize" it. To assist with the synchronization process, investigators across the two studies used several strategies; consequently, the synchronization process resulted in more of an art than a science. Specifically, investigators utilized audiovisual cues from the video recording, the event-mark buttons on the sensors, and the accelerometer data to assist with synchronization. Finally, the internal clock of the sensor was compared to the time the session started, indicated by the video data (synchronization strategies are highlighted in the considerations for practice section 3.5.2.3). Unfortunately, two main challenges arose that made the synchronization process between the EDA data and video more difficult. One, children were very interested in exploring the sensors and would often press the event-mark buttons on the sensors, and two, the internal time of the sensors was not exactly synchronized with real-time. Initially, investigators in the SL-EDA study used ELAN to assist with synchronization of the visual and the EDA data but the software proved unreliable and would often "freeze" and "crash". In contrast, investigators in the OT-EDA study were able to use BEDA to

assist with the synchronization process successfully, a program that was specifically designed for this purpose by one of the investigators.

3.4.2.3 Data preprocessing/analysis. The third area of technical challenge that emerged in relation to use of EDA data was related to data preprocessing and analysis. Even though EDA has been used widely for over 130 years, reliable software to analyze *in situ* continuous EDA data is still in development and with little in the way of standardized procedures (cf. see Boucsein (1993) 2012 for the most comprehensive review of psychological applications, mechanisms and methodology of EDA). Even without ELAN's freezing and crashing episodes, the software was not able to do what the investigators across studies wanted to do such as comparing many sessions at the same time. Initially, investigators across the two studies ended up having to use multiple software programs for multiple purposes, such as using one program to synchronize the video and EDA data and another to crop the EDA data. Even the investigator with prior signal processing experience had difficulty pre-processing and analyzing the data with the available software programs (e.g., Ledalab). When asked about the software available to analyze the EDA data, the OT-EDA investigator acknowledged she "had some difficulties importing the signal" and difficulties in general, "that's why" BEDA was developed (Kim et al. 2013), to assist with visualization, synchronization and analysis of the data.

3.5 Considerations for Practice

3.5.1 Behavioral considerations: Desensitization techniques & additional behavioral strategies. Being exposed to a novel object by a new person in an unfamiliar environment can be quite difficult for many children, particularly children with neurodevelopmental disabilities who may be particularly sensitive to certain stimuli and have difficulty understanding and being understood by others. Desensitization procedures have been successfully used to reduce anxiety in children when exposed to novel technology. To illustrate, Barnea-Goraly et al. (2014) successfully used a brief behavioral training with children ages 4 -10 during MRI scans as a means to eliminate the use of sedation to acquire motion-free high-quality images.

Behavioral findings across the SL and OT EDA studies suggest younger children (2-4 years) with neurodevelopmental communication impairments require a period of desensitization to the biosensors before wearing the biosensors. Based on our data, we recommend that investigators using similar biosensors allow at least 4 desensitization sessions, prior to data collection sessions,

for children below the age of 5. Although older children may require less time and attention to the desensitization process, we recommend planning at least one desensitization session when working with children, particularly those with marked impairments or differences. Consequently, we offer here the explicit desensitization techniques we used to acclimate children to the sensors (see Table 3.3 for a summary of the behavioral considerations). Desensitization procedures have been defined as "pairing of either graduated imagined or graduated external stimuli with either relaxation or other responses competitive with anxiety" (Hatzenbuehler and Schroeder 1978) as a means to decrease the anxiety associated with the stimuli, and are common in disciplines such as special education and psychology. Consistent with systematic desensitization procedures (Hatzenbuehler and Schroeder 1978, Ollendick and Cerny (1981) 2013), our desensitization procedures focused on presentation of stimuli (i.e., the sensors) in a graduated hierarchy, gradual pairing of the positive valence responses with the sensors, and in creating an environment conducive to positive valence responses.

3.5.1.1 Presentation of biosensors in a graduated hierarchy. To desensitize children to wearing the sensors, we recommend presenting the sensors in a graduated hierarchy, progressively increasing the child's exposure to them. Prior to direct expectations to wear the sensors (i.e., in-home exposure), caregivers were encouraged to read to the child a social story we prepared to explain the process of wearing the biosensors. Consistent with the use of social stories, often used to familiarize children with developmental disabilities to behavioral expectations in new situations and to reduce challenging behaviors, the stories were written from a first-person perspective and included photographs as visual supports (Gray and Garand 1993, Moudry Quilty 2007, Swaggart et al. 1995); see Appendix A for our specific example. In addition to being asked to share the story with their child, caregivers were encouraged to fit their children with bracelets/watches around ankles/wrists at home –based on intended sensor placement- for increasing periods of time to provide their child with the opportunity to experience sensations similar to those they would experience when wearing the sensors. They were also encouraged to bring the bracelets/watches to the initial sessions for visual support when drawing comparisons with the sensors.

Desensitization techniques employed during direct exposure to the sensors (i.e., utilized during the observation, initial assessment and treatment sessions) included rereading the prepared social story, providing reassurance, and additional strategies to gradually familiarize children with the sensors. This included incorporating a picture of the sensors in the child's visual schedule of the

Table 3.3. Behavioral Considerations for Short Term *In Situ* Skin Conductance Collection & Analyses using Wearable Biosensors with Children with Neurodevelopmental Communication Impairments

General Considerations
– Use desensitization procedures when exposing children to novel technology. At least one desensitization period/session is recommended for children over five years; younger children are likely to require more (≥4).
– Out of sight, out of mind: place sensors out of sight, preferable on the ankles, and wrap each sensor with a brace. Each sensor should have a snug but not a tight fit as it should fit the child comfortably.
Present Biosensors in a Graduated Hierarchy, Progressively Increase Exposure to the Biosensors
In- home Exposure:
– Developing and reading a social story (see Appendix A)
– Fit the child with bracelets/watches (or sensors) around ankles/wrists at home for increasing periods of time
On Site Exposure to Biosensors:
– Encourage the caregiver to actively participate in the desensitization process
– Incorporate a picture of the sensors in the child's visual schedule of the session's activities
– Reread social story at the onset of sessions
– Draw comparisons between bracelets/watches used at home and sensors (e.g. “It is just like Papa's watch!”)
– Encourage child to explore the sensors (e.g. make note of the flashing light) and provide reassurance
– Place the sensors first on the therapist/researcher, then on the caregiver and finally on the child. If the child is apprehensive, move the biosensors incrementally closer to the child and encourage the child to wear the biosensors for a specified period of time (e.g. 2 seconds. Count the seconds out loud and/or use a timer for visual support (e.g. hourglass, iPad app)).
Gradually Pair Positive Valence Responses with the Biosensors
– Introduce the sensors to the child when the child is in a positive affective state (e.g. excited, attentive)
– Decorate the sensors e.g., with stickers or other decorations to make them more appealing to the child
– Refocus the child's attention to a preferred activity (i.e., one that promotes positive affective states)
– Use reinforcement procedures such as if-then statements, verbal and non-verbal praise for attempts, and game playing
Create an Environment Conducive to Positive Valence Responses
– Build rapport and trust with the child
– Tailor the strategies & reinforcers used to the child's individual needs and preferences
– Introduce new activities and reinforcers for novelty as needed, and make the sessions as engaging and enjoyable as possible!

session's activities (cf. Mirenda and Brown 2009), introducing the biosensors as watches/bracelets[15], and providing a period in which children were encouraged to explore the biosensors themselves, gradually increasing exposure to the sensors consistent with systematic desensitization literature. Specifically, the investigator might draw the child's attention to different aspects of the sensors, such as the flashing light. If the child is hesitant to touch the biosensors at all, an investigator might begin by manipulating them themselves and subsequently putting them on the caregiver's wrist/ankle during initial presentation. The presence of the child's caregivers during the initial sessions also served to increase comfort. Other suggestions might include moving the biosensors incrementally closer to the child's wrist or ankle or encouraging the child to wear the biosensors for a specified short period of time (e.g. 2 seconds) and count the seconds out loud or use a timer (e.g. hourglass, iPad app) for visual support, which can be progressively increased in time.

3.5.1.2 Gradual pairing of positive valence responses with sensors. As a means to gradually pair the sensors with positive valence responses, investigators introduced the sensors to the child when the child appeared to be in a positive affective state (e.g. interested, excited, proud, and attentive vs. negative affective states such as upset, scared, and nervous). The biosensor itself can also be decorated with stickers or other images and colors to make it more familiar and appealing to the child. For example, we decorated Dora's biosensors with images of Dora from "Dora the Explorer" after learning she highly enjoyed watching the show at home during the initial assessment session prior to the start of therapy. Similarly, refocusing attention from the biosensors to a favorable activity that elicits responses competitive with anxiety, such as excitement or relaxation, might be helpful (e.g. "Let's go bounce on the ball!", "time to play with the iPad", "time to read your favorite book"). Finally, reinforcement procedures in which children were rewarded by their parents or clinicians after wearing the sensors were also used. For example, Pan was taken to the pool, a favored activity for him, after the sessions (e.g. "**If** you wear the sensors, **then** we can go to the pool after speech" would be an example of delayed reinforcement). Verbal and non-verbal praise for attempts and game playing were also used to provide immediate positive reinforcement (e.g. sensors are introduced when child appears to be in a positive affective state →

[15] Excerpt indicating how clinicians across the studies introduced the sensors to the children: "These are bracelets/watches [show child sensors] that feel and look very similar to your bracelets/watches [point to child's bracelets/watches they wore at home]. They will give us information about your feelings. For example, they will tell us whether you are excited or relaxed."

child wears sensors → adults clap, smile, cheer, provide specific verbal praise, and provide child with tangible reinforcer such as the iPad).

3.5.1.3 Creating an environment conducive to positive valence responses. Finally, we focused in creating an environment conducive to positive valence responses. For example, caregivers were asked to provide information regarding their child's preferred activities and objects/toys, individual child's preferences were informally assessed throughout the intervention, and new activities and objects/toys were introduced for novelty. By the end of the study, children and the investigators grew mutually fond of one another. This was evidenced by children expressing their desire to continue with therapy after being told it was clean-up time, and by caregiver's unsolicited report of the children inquiring about their clinicians while at home.

Whatever specific procedures are employed, it is important to anticipate the need for individualized responses. Children are likely to respond differentially to the same strategy. For example, whereas gradually increasing contact with the biosensors worked wonderfully for Angelo, this same strategy was not effective for Heidi.

3.5.2 Technical considerations.

3.5.2.1 Factors that may influence EDA. *In situ* or ambulatory skin conductance assessment has made major advances over the last decade but continues to require close monitoring of the multiple factors that may influence the data including ambient temperature, time of day, physical activity, and individual differences (e.g., race, age, sex). In addition to requiring careful monitoring, decreased ability to control such factors requires careful selection of the preprocessing and analysis methods and careful interpretation of the results. Consequently, few *in situ* EDA studies have been published to date. Of those reported, some offer only superficial assessments of EDA and/or had to discard a high percentage of the data. For example, Miller, Coll, and Schoen (2007) monitored EDA while administering a series of sensory stimuli in 24 children with sensory modulation disorders and had to discard 54% of the EDA data and provided few details regarding EDA data analysis. Doberenz et al. (2011) assessed multiple factors known to influence SC during a 24-hour period *in situ* and concluded that although feasible, some measures need to be corrected for the influence of confounding variables. Other reports on problems of long-term *in situ* EDA recording have found the conducting medium or gel to significantly influence EDA recording, with a hydrating medium leading to fewer and smaller SCRs compared to a non-hydrating medium

(Turpin, Shine, and Lader 1983), and have even concluded ambulatory recordings to lack both reliability and validity after 24 hours, although this particular study had low power and did not conduct statistical comparisons (Boucsein, Schaefer, and Sommer 2001). Given difficulties with wet electrode use in long-term ambulatory SC assessment, dry electrodes are also available.

Although research examining short-term ambulatory SC poses fewer challenges compared with long-term ambulatory SC recording (e.g., electrode deterioration is less of a concern), it poses challenges nonetheless. Overall, it is important to note skin conductance tends to vary widely across subjects (i.e. high inter-individual variance) and is more stable within subjects (Dawson, Schell, and Filion (2000) 2007). It is also important to highlight that the same factors (e.g. physical activity) may influence between-subject comparisons and within-subject comparisons differentially (Doberenz et al. 2011). Researchers interested in assessing skin conductance need at least a basic understanding in signal processing, time series data, and in the physiological basis of electrodermal activity and are referred to Boucsein's ((1993) 2012) electrodermal activity book for a more thorough review and Dawson and colleagues' ((2000) 2007) chapter on electrodermal activity in the handbook of psychophysiology for a shorter review. The remainder of this section will provide recommendations and references based on the factors likely to influence short term *in situ* monitoring of EDA including ambient temperature, time of day, physical activity, and individual differences but it is by no means inclusive, and its intended purpose is to make readers aware of some of the factors needing special consideration. Here we offer explicit considerations to address the technical challenges encountered (see Table 3.4 for a summary of the technical considerations).

3.5.2.1.1 Ambient temperature & time of day. In regards to ambient temperature and time of day, recording in a temperature-controlled setting during the same time of day throughout the study is advised. Ambient temperature is positively correlated with the frequency of SCRs between subjects (Doberenz et al. 2011, Turpin, Shine, and Lader 1983) and is significantly positively correlated with a variety of EDA measures within subjects (Doberenz et al. 2011) and EDA has been reported to be lowest in the morning (Miro, Cano-Lozano, and Buela-Casal 2002). Related to this, hydration of the skin (both endogenous and exogenous) may also influence EDA highlighting the importance of the medium of conduction used. Electrode gels containing either KCl or NaCl are recommended during short-term EDA assessment to avoid variations in EDA, and improve signal acquisition, as both of these appear as salts in the stratum corneum

Table 3.4. Technical Considerations for Short Term *In Situ* Skin Conductance Collection & Analyses using Wearable Biosensors with Children with Neurodevelopmental Communication Impairments

General Considerations

- Researchers are recommended to become familiar with the following electrodermal activity resources: Boucsein ((1993) 2012), Dawson et al.((2000) 2007)

Control for Factors that may Influence EDA, including:

- Ambient temperature and time of day (i.e. record data in a temperature-controlled setting (usually 22-24 °C), during the same time of day)
- Physical activity (e.g. walking, running, bouncing, jumping and/or other repetitive motor movements such as hand flapping)
- Individual differences (e.g. age, sex, race, body mass index and sweat gland density)
- Use of medications and psychoactive drugs (e.g. caffeine)

Sensor Functioning

- Collect data using two sensors (one on each side of the body), and ensure at least one other sensor is available for data collection in case one malfunctions.
- Use internal storage over bluetooth: bluetooth was often disconnected and data collected resulted in a drift
- Given data across sensors are comparable: Use dominant wrist/ankle to yield higher amplitude. Use nondominant site to decrease potential movement artifacts.
- Use KCL or NaCL electrode gel as the conductive medium between electrode and the skin. Implement a 15 minute biosensor acclimation period for the skin-electrolyte interface to stabilize.
- Back up data after every session & delete files stored in the sensor

Synchronization

- Synchronize the sensor time with real-time before each treatment session
- Verbally narrate relevant events (e.g. sensor A is on the participant's right ankle and first flashed at 8:03:05 am)
- Use the event-mark button for relevant events (e.g. start and end of treatment sessions)
- After placing sensors on child & marking an event, move sensors back & forth in front of the camera to synchronize accelerometer & video data
- Use a visual analytic tool to visualize and synchronize the data (e.g. BEDA).

Data Preprocessing & Analysis

- Analyze both the phasic and tonic components of EDA (for deconvolution approach see Benedek and Kaernbach 2010b)
- Use a visual analytic tool to assist with data preprocessing & analysis (e.g. BEDA)
- Normalization across subjects is needed for between-subject comparisons (e.g. Hernandez et al. 2014 normalized the range between 0-1)

Recommended Thresholds:

- Exclusion of SCRs if greater than 1.0 µS (Dawson et al. (2000) 2007; with Doberenz et al. 2011 recommending a .5 µS threshold) as this likely represents movement artifacts or loose electrodes
- Exclusion of SC data if values are below .5 µS (Doberenz et al. 2011) given the typical SCL range is between 2 µs and 20 µs (Dawson et al. (2000) 2007))
- Minimum amplitude threshold of .01 µS although criterion is ultimately dependent on the resolution of the recording (Boucsein (1993) 2012, 156-157, Dawson, Schell, and Filion (2000) 2007, Hernandez et al. 2014)
- Minimum distance between NS.SCRs of 1 second (Hernandez et al. 2014)

(readers are referred to Boucsein (1993) 2012, 117 for additional details). In order for the skin-electrolyte interface to stabilize, a biosensor acclimation period of approximately 10-15 minutes should be implemented.

3.5.2.1.2 Physical activity. Physical activity may influence SC via thermoregulation. However, associations between motor movement within daily activities in humans and skin conductance is not yet well elucidated, especially in children. Turpin, Shine, and Lader (1983) assessed long-term ambulatory SC recorded from the fingers in a group of 12 adults, and found no between-subjects or mean within-subject significant correlations between arm movements and EDA. On the other hand, within-subject comparisons in a group of 48 healthy adults who wore ambulatory SC devices on their fingers for a 24 hour period, showed significant positive relationships between physical activity and EDA with the exception of amplitude of NS.SCRs which decreased with increased physical activity (Doberenz et al. 2011). As such, although corrections for the effects of confounding variables such as physical activity may not be necessary at the between-subject level, corrections are recommended at the within-subject level. For the SL-EDA study, even though physical activity was somewhat constant within subjects across sessions, physical activity such as walking, running, bouncing, jumping and/or engaging in repetitive motor mannerisms led to movement of the electrodes and/or to loose electrodes thereby increasing artifacts in the data. As previously discussed securing the electrodes with medical tape, a wrist or ankle wrap is advised.

3.5.2.1.3 Individual differences. Given that EDA may vary with differences in race, age, sex, body mass index, sweat gland density, and use of medications and psychoactive substances (Doberenz et al. 2011), controlling for such factors is advised and normalization across subjects is needed for between-subject comparisons (e.g. Hernandez et al. 2014 normalized the range of values to be between zero and one). For additional information regarding race/ethnicity differences in SC in particular, see Wesley and Maibach (2003). For a detailed review of the effects of and interactions with individual differences more broadly see Boucsein ((1993) 2012).

3.5.2.2 Sensor functioning. Investigators are encouraged to collect data using two sensors (one on each side of the body) given potential EDA asymmetry (Picard, Fedor, and Ayzenberg 2016), and to have at least one other sensor readily available to use for data collection in case one of the sensors malfunctions. The sensors should be calibrated to be the same, and data collected across the two sensors should be visually inspected. Given the challenges experienced with the Bluetooth

device, we recommend using the internal storage for data collection and analysis. Additionally, depending on specific protocol to follow (e.g. handwriting) or depending on participant (e.g. child's basal SC level is low), as long as data across sensors are comparable, whether data are analyzed for the dominant or nondominant wrist/ankle may vary. If interested in yielding a higher amplitude, analyze data from the dominant wrist/ankle (see Román et al. 1989). If interested in decreasing potential movement artifacts, analyze data from the nondominant wrist/ankle. Finally, to avoid exceeding the threshold of directories stored in the device, we recommend backing up the data after every session and deleting the files in the sensor.

3.5.2.3 Synchronization. We would like to remind researchers to synchronize the internal sensor-time with real-time prior to the onset of each treatment session. Moreover, additional strategies to assist with synchronizing audio-video recorded data with physiological data include: a) verbally narrate relevant events for the recording (e.g., "sensor 'A' is on the participant's right ankle and first flashed at 8:03:05 am"); b) press the event-mark buttons on the sensors for relevant events (e.g., at the beginning and end of the sessions); and c), after placing the sensors on the child and marking an event at the beginning and end of the session, move the sensor back and forth for a few seconds in front of the camera to synchronize the accelerometer data with the video data. The latter was the preferred synchronization strategy by the researchers. Finally, we advise researchers to use a visual analytic tool to visualize, synchronize and analyze the data such as BEDA.

3.5.2.4 Data preprocessing & analysis. Analyzing both the phasic and tonic components of EDA is recommended as these can aid in the interpretation of overall skin conductance results (without decomposition) and aid in the understanding of the physiological/neurological underpinnings of the various EDA measures. Researchers interested in exploring the deconvolution approach are referred to Benedek and Kaernbach (2010b). Although future studies should address what the optimal thresholds for preprocessing and analysis of short-term *in situ* SC are in order to establish a recommended set of thresholds, based on the current literature the following are recommended: a) exclusion of SCRs if greater than 1.0 μS (Dawson, Schell, and Filion (2000) 2007, with Doberenz et al. 2011 recommending a .5 μS threshold) as this likely represents movement artifacts or loose electrodes, b) exclusion of SC data if values are below .5 μS (Doberenz et al. 2011) given the typical SCL range is between 2 μS and 20 μS (Dawson, Schell, and Filion (2000) 2007), c) a minimum amplitude threshold of .01 although criterion is ultimately

dependent on the resolution of the recording (Boucsein (1993) 2012, 156-157, Dawson, Schell, and Filion (2000) 2007, Hernandez et al. 2014), and d) a minimum distance between NS.SCRs of 1 second (Hernandez et al. 2014).

3.6 Conclusions & Future Directions

Skin conductance has largely been monitored in highly controlled experimental conditions and more recently has been successfully monitored *in situ* in individuals with generalized tonic-clonic seizures who present with significantly larger and more frequent skin conductance responses, but less so to monitor more general variation during daily activities. The present work is novel in presenting behavioral and technical methodological considerations, derived from novel post hoc analyses, when monitoring EDA in children with neurodevelopmental communication impairments based on the results of two studies examining continuous short-term EDA *in situ* during speech-language and occupational therapy.

The SL- EDA study monitored skin conductance in children at the single-word developmental stage undergoing speech-language treatment to increase their multisyllabic productions and examined associations between skin conductance and emotional valence. The OT-EDA study monitored skin conductance in children with sensory processing difficulties undergoing occupational therapy and examined associations between skin conductance, use of a pressure vest, type of instruction, academic engagement and challenging behaviors. Younger children (ages 2-4) children wore the sensors for approximately 62% of the sessions, whereas older children (ages 9 and 11) wore the sensors for all of the sessions. Unfortunately, due to a combination of behavioral and technical challenges a large percentage of data were discarded (i.e., 29%) and is often discarded when monitoring EDA in children and/or in applied settings (e.g., 31% of the data were discarded in Hernandez et al. 2014, and 54% in Miller, Coll, and Schoen 2007). We presented behavioral and technical considerations including the use of desensitization techniques and recommended thresholds based on the current literature for preprocessing and analysis of EDA data.

Moving forward, research should focus on how to improve signal acquisition, due to both behavioral and technical challenges, to decrease overall percentage of data being discarded, on how to simplify the process of synchronization across multiple data sources within the field of accessible computing, and on establishing a set of recommended thresholds for acquisition of short

term *in situ* SC data. In sum, although the present work shows it is feasible to record EDA *in situ,* it highlights many of the challenges of monitoring EDA in applied settings, or in other words in uncontrolled environments. It is of particular importance that we recognize and address these challenges before commercializing the use of biosensors that measure SC to children with neurodevelopmental communication impairments. Ultimately, we aim to aid in the development and improvement of automated noninvasive unobtrusive and easy to interpret tools for measuring levels of emotional arousal via skin conductance, the only noninvasive autonomic nervous system measure innervated solely by and consequently most representative of sympathetic nervous system activity. Although not yet ready to be adopted within clinical practice or within the homes of children with neurodevelopmental communication impairments, EDA technological advances offer a unique opportunity to assess changes in emotional arousal that may help guide teaching opportunities for children with neurodevelopmental communication impairments.

Acknowledgments: The SL-EDA intervention study was funded by Autism Speaks (#5744); the OT-EDA study was funded by the Computation Behavioral Science: Modeling, Analysis, and Visualization of Social and Communicative Behavior, NSF Expeditions Grant (#1029035) and The Autism Program of Illinois; MAB was funded by the Neuroscience Program Fellowship, the Goldstick Family Doctoral Fellowship for the study of communication disorders, and the Child Language & Literacy Laboratory. We are appreciative of all participating families who made this study possible. We acknowledge the SL-EDA study investigators and thank participating members of the Child Language & Literacy Lab, specifically: Ellen Bogue, Paulina Mitra, Lauren Ragins, Monique Kammo, Emily Zimmerman, Katherine Dejmek, Kori Zorina, Shanthi Sivasankaran, and Carley Serena. Finally, we acknowledge the OT-EDA study investigators: Melinda Snodgrass, Jim Halle, Mindy Borden, Mary Pietrowicz, and Hedda Meadan-Kaplansky.

References

Affectiva. 2012. "Q Sensor 2.0." Last Accessed December, 2012. http://www.affectiva.com/q-sensor/.

Affectiva. 2013. "Q User Manual." Last Modified March 29, 2013. http://qsensor-support.affectiva.com/customer/en/portal/articles/1539375-q-user-manual.

APA, (American Psychiatric Association). 2013. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. Arlington, Virginia: American Psychiatric Publishing.

Aparicio Betancourt, Mariana, Laura DeThorne, and Karrie Karahalios. 2013. "Facilitating Multisyllabic Productions & Assessing Sympathetic Arousal in Children with Developmental Disorders." Poster presented at the Society for Neuroscience Conference, San Diego, California, November 9-13. http://hdl.handle.net/2142/78186.

ASHA, (American Speech-Language-Hearing Association). 2014. "New Pediatrics Study Reveals Dramatic Increase in Speech Problems in Children Over Past Decade." Last Modified August 21. http://www.asha.org/About/news/Press-Releases/2014/New-Pediatrics-Study-Reveals-Dramatic-Increase-in-Speech-Problems-in-Children-Over-Past-Decade.htm.

Autism&Oughtisms. 2011. "“Has Autism” versus “Is Autistic”; A Muddled Debate." Autism & Oughtisms, Last Modified July 25. https://autismandoughtisms.wordpress.com/2011/07/25/has-autism-versus-is-autistic-a-muddled-debate/.

Barnea-Goraly, Naama, Stuart A. Weinzimer, Katrina J. Ruedy, Nelly Mauras, Roy W. Beck, Matt J. Marzelli, Paul K. Mazaika, Tandy Aye, Neil H. White, Eva Tsalikian, Larry Fox, Craig Kollman, Peiyao Cheng, and Allan L. Reiss. 2014. "High Success Rates of Sedation-free Brain MRI Scanning in Young Children using Simple Subject Preparation Protocols with and without a Commercial Mock Scanner--the Diabetes Research in Children Network (DirecNet) Experience." *Pediatric Radiology* 44 (2):181-6. doi: 10.1007/s00247-013-2798-7.

Barton, Erin E., Brian Reichow, Alana Schnitz, Isaac C. Smith, and Daniel Sherlock. 2015. "A Systematic Review of Sensory-based Treatments for Children with Disabilities." *Research in Developmental Disabilities* 37:64-80. doi: 10.1016/j.ridd.2014.11.006.

Benedek, Mathias, and Christian Kaernbach. 2010a. "A Continuous Measure of Phasic Electrodermal Activity." *Journal of Neuroscience Methods* 190 (1-5):80-91. doi: 10.1016/j.jneumeth.2010.04.028.

Benedek, Mathias, and Christian Kaernbach. 2010b. "Decomposition of Skin Conductance Data by means of Nonnegative Deconvolution." *Psychophysiology* 47 (4):647-58. doi: 10.1111/j.1469-8986.2009.00972.x.

Berez, Andrea. L. 2007. "Technology Review: EUDICO Linguistic Annotator (ELAN)." *Language Documentation & Conservation* 1 (2):283-289. http://scholarspace.manoa.hawaii.edu/html/10125/1718/berez.pdf.

Boucsein, Wolfram. (1993) 2012. *Electrodermal Activity*. 2nd ed. Reprint, New York, New York: Springer Science & Business Media.

Boucsein, Wolfram, Florian Schaefer, and Thorsten Sommer. 2001. "Electrodermal Long-term Monitoring in Everyday Life." In *Progress in Ambulatory Assessment: Computer-assisted Psychological and Psychophysiological Methods in Monitoring Field Studies*, edited by Jochen Fahrenberg and Michael Myrtek, 549-560. Cambridge, Massachusetts: Hogrefe & Huber Publishers (quoted in Doberenz et al. 2011).

Braun, Virginia, and Victoria Clarke. 2006. "Using Thematic Analysis in Psychology." *Qualitative Research in Psychology* 3 (2):77-101. doi: 10.1191/1478088706qp063oa.

Brown, Rachael, Cheree James, Luke Henderson, and Vaughan Macefield. 2012. "Autonomic Markers of Emotional Processing: Skin Sympathetic Nerve Activity in Humans during Exposure to Emotionally Charged Images." *Frontiers in Physiology* 3 (394). doi: 10.3389/fphys.2012.00394.

Brugman, Hennie, and Albert Russel. 2004. "Annotating Multi-media/ Multi-modal Resources with ELAN." Proceedings of Language Resources and Evaluation Fourth International Conference (LREC).

Critchley, Hugo D. 2002. "Electrodermal Responses: What Happens in the Brain." *The Neuroscientist* 8 (2):132-142. doi: 10.1177/107385840200800209.

Curtis, Sophie. 2015. "Biometric Wristbands Predict Outbursts in People with Autism." The Telegraph, Last Modified September 16. http://www.telegraph.co.uk/technology/news/11866862/Biometric-wristbands-predict-outbursts-in-people-with-autism.html.

Dawson, Michael E., Anne M. Schell, and Diane L. Filion. (2000) 2007. "The Electrodermal System." In *Handbook of Psychophysiology*, edited by John T. Cacioppo, Louis G. Tassinary and Gary Berntson, 200-223. Reprint, New York, New York: Cambridge University Press.

DeThorne, Laura, Mariana Aparicio Betancourt, Karrie Karahalios, Jim Halle, and Ellen Bogue. 2015. "Visualizing Syllables: Real-time Computerized Feedback within a Speech-language Intervention." *Journal of Autism and Developmental Disorders* 45 (11):3756-63. doi: 10.1007/s10803-014-2274-8.

Doberenz, Sigrun, Walton T. Roth, Eileen Wollburg, Nina I. Maslowski, and Sunyoung Kim. 2011. "Methodological Considerations in Ambulatory Skin Conductance Monitoring." *International Journal of Psychophysiology* 80 (2):87-95. doi: 10.1016/j.ijpsycho.2011.02.002.

Duncan, Stuart. 2011. "The Last Word on "Person First" Language." Thinking Person's Guide to Autism, Last Modified July 20. http://www.thinkingautismguide.com/2011/07/last-word-on-person-first-language.html

el Kaliouby, Rana, and Matthew S. Goodwin. 2008. "iSET: Interactive Social-emotional Toolkit for Autism Spectrum Disorder." Proceedings of the 7th International Conference on Interaction Design and Children, Workshop on Special Needs, Chicago, IL, June 11-13.

Empatica. 2016a. "Embrace Watch Epilepsy Monitor." Last Accessed March, 2016. https://www.empatica.com/embrace-watch-epilepsy-monitor.

Empatica. 2016b. "Embrace Watch Technology." Last Accessed March, 2016. https://www.empatica.com/embrace-watch-technology.

Figner, Bernd, and Ryan O. Murphy. 2011. "Using Skin Conductance in Judgment and Decision Making Research." In *A Handbook of Process Tracing Methods for Decision Research*, edited by Michael Schulte-Mecklenbeck, Anton Kuehberger and Rob Ranyard, 163-184. New York, New York: Psychology Press.

Fleischmann, Arthur, and Carly Fleischmann. 2012. *Carly's Voice: Breaking Through Autism*. New York, NY: Simon and Schuster.

Fletcher, Richard R., Kelly Dobson, Matthew S. Goodwin, Hoda Eydgahi, Oliver Wilder-Smith, David Fernholz, Yuta Kuboyama, Elliot B. Hedman, Ming-Zher Poh, and Rosalind W. Picard. 2010. "iCalm: Wearable Sensor and Network Architecture for Wirelessly Communicating and Logging Autonomic Activity." *IEEE Transactions on Information Technology in Biomedicine* 14 (2):215-223. doi: 10.1109/TITB.2009.2038692.

Gainey, Summer. 2013. "Increasing Communicative Mand Topographies: An Evaluation of the Use of a Lag Reinforcement Schedule Embedded in a Functional Communication Training Treatment Package." PhD diss., University of Texas at Austin.

Grace, Elizabeth J. 2015. "Socio-Sensory Distress is Harsh, but Navigable." In *The Real Experts: Readings for Parents of Autistic Children*, edited by Michelle Sutton, 74-77. Fort Worth, Texas: Autonomous Press.

Grandin, Temple. (1992) 2013. "An Inside View of Autism." In *High-functioning Individuals with Autism*, edited by Eric Schopler and Gary B. Mesibov, 105-126. Reprint, New York, New York: Springer.

Gray, Carol A., and Joy D. Garand. 1993. "Social Stories: Improving Responses of Students with Autism with Accurate Social Information." *Focus on Autistic Behavior* 8 (1):1-10. doi: 10.1177/108835769300800101.

Hagbarth, Karl E., Rolf G. Hallin, Anders Hongell, H. Erik Torebjörk, and B. Gunnar. Wallin. 1972. "General Characteristics of Sympathetic Activity in Human Skin Nerves." *Acta Physiologica Scandinavica* 84 (2):164-176. doi: 10.1111/j.1748-1716.1972.tb05167.x.

Halle, James W., Michaelene M. Ostrosky, and Mary Louise Hemmeter. 2006. "Functional Communication Training: A Strategy for Ameliorating Challenging Behavior." In *Treatment of Language Disorders in Children*, edited by R.J. McCauley and M.E. Fey, 509–548. Baltimore, Maryland: Paul H. Brookes.

Hatzenbuehler, Linda C., and Harold E. Schroeder. 1978. "Desensitization Procedures in the Treatment of Childhood Disorders." *Psychological Bulletin* 85 (4):831-844. doi: 10.1037/0033-2909.85.4.831.

Healey, Jennifer A., and Rosalind W. Picard. 2005. "Detecting Stress during Real-world Driving Tasks using Physiological Sensors." *IEEE Transactions on Intelligent Transportation Systems* 6 (2):156-166. doi: 10.1109/TITS.2005.848368.

Hedman, Elliott B. 2010. "*In-situ* Measurement of Electrodermal Activity during Occupational Therapy." Master's thesis, Architecture and Planning, Program in Media Arts and Sciences, Massachusetts Institute of Technology.

Hemmeter, Mary Louise, Michaelene Ostrosky, and Lise Fox. 2006. "Social and Emotional Foundations for Early Learning: A Conceptual Model for Intervention." *School Psychology Review* 35 (4):583-601.

Hernandez, Javier, Ivan Riobo, Agata Rozga, Gregory D. Abowd, and Rosalind W. Picard. 2014. "Using Electrodermal Activity to Recognize ease of Engagement in Children during Social Interactions." Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing, Seattle, Washington.

Hillary, Alyssa. 2015. "Autism and Language." In *The Real Experts: Readings for Parents of Autistic Children* edited by Michelle Sutton, 35-36. Fort Worth, Texas: Autonomous Press.

Hodgetts, Sandra, Joyce Magill-Evans, and John E. Misiaszek. 2011. "Weighted Vests, Stereotyped Behaviors and Arousal in Children with Autism." *Journal of Autism and Developmental Disorders* 41 (6):805-14. doi: 10.1007/s10803-010-1104-x.

Hutchins, Tiffany L., and Patricia A. Prelock. 2014. "Using Communication to Reduce Challenging Behaviors in Individuals with Autism Spectrum Disorders and Intellectual Disability." *Child and Adolescent Psychiatric Clinics of North America* 23 (1):41-55. doi: 10.1016/j.chc.2013.07.003.

Jaques, Natasha, Sara Taylor, Asaph Azaria, Asma Ghandeharioun, Akane Sano, and Rosalind Picard. 2015. "Predicting Students' Happiness from Physiology, Phone, Mobility, and Behavioral Data." International Conference on Affective Computing and Intelligent Interaction (ACII), September 21-24.

Kim, Jennifer, Melinda Snodgrass, Mary Pietrowicz, Karrie Karahalios, and Jim Halle. 2013. "BEDA: Visual Analytics for Behavioral and Physiological Data." Proceedings of the Workshop on Visual Analytics in Health Care, Washington DC, MD, November 16th.

Kraft, Jessica C. 2015. "A New Smartwatch that Detects Seizures and Emotional Stress." KQED Science, Last Modified June 4. http://ww2.kqed.org/futureofyou/2015/06/04/a-new-smartwatch-that-detects-seizures-and-emotional-stress/.

Kunimoto, Masanari, Klaus Kirnö, Mikael Elam, Tomas Karlsson, and B. Gunnar Wallin. 1992. "Neuro-effector Characteristics of Sweat Glands in the Human Hand Activated by Irregular Stimuli." *Acta Physiologica Scandinavica* 146 (2):261-269. doi: 10.1111/j.1748-1716.1992.tb09415.x.

Kunimoto, Masanari, Klaus Kirnö, Mikael Elam, and B. Gunnar Wallin. 1991. "Neuroeffector Characteristics of Sweat Glands in the Human Hand Activated by Regular Neural Stimuli." *The Journal of Physiology* 442:391-411. doi: 10.1113/jphysiol.1991.sp018799.

Kylliainen, Anneli, and Jari K. Hietanen. 2006. "Skin Conductance Responses to another Person's Gaze in Children with Autism." *Journal of Autism and Developmental Disorders* 36 (4):517-25. doi: 10.1007/s10803-006-0091-4.

Lang, Russell, Mark O'Reilly, Olive Healy, Mandy Rispoli, Helena Lydon, William Streusand, Tonya Davis, Soyeon Kang, Jeff Sigafoos, Giulio Lancioni, Robert Didden, and Sanne Giesbers. 2012. "Sensory Integration Therapy for Autism Spectrum Disorders: A Systematic Review." *Research in Autism Spectrum Disorders* 6 (3):1004-1018. doi: http://dx.doi.org/10.1016/j.rasd.2012.01.006.

Law, James, James Boyle, Frances Harris, Avril Harkness, and Chad Nye. 2000. "Prevalence and Natural History of Primary Speech and Language Delay: Findings from a Systematic Review of the Literature." *International Journal of Language and Communication Disorders* 35 (2): 165-188.

Lee, Chia-Hsun Jackie, Rob Morris, Matthew Goodwin, and Rosalind Picard. 2008. "Lessons Learned from a Pilot Study Quantifying Face Contact and Skin Conductance in Teens with Asperger Syndrome." CHI '08 Extended Abstracts on Human Factors in Computing Systems, Florence, Italy, April 5-10.

Madsen, Miriam A. 2010. "Developing an Interactive Social-Emotional Toolkit for Autism Spectrum Disorders." Master's thesis, Electrical Engineering and Computer Science, Massachusetts Institute of Technology.

Mangina, Constantine A., and J. Helen Beuzeron-Mangina. 1996. "Direct Electrical Stimulation of Specific Human Brain Structures and Bilateral Electrodermal Activity." *International Journal of Psychophysiology* 22 (1-2):1-8. doi: 10.1016/0167-8760(96)00022-0.

Miller, Lucy J., Joseph R. Coll, and Sarah A. Schoen. 2007. "A Randomized Controlled Pilot Study of the Effectiveness of Occupational Therapy for Children with Sensory Modulation Disorder." *American Journal of Occupational Therapy* 61 (2):228-38. doi: 10.5014/ajot.61.2.228.

Ming, Xue, Peter O. Julu, Michael Brimacombe, Susan Connor, and Mary L. Daniels. 2005. "Reduced Cardiac Parasympathetic Activity in Children with Autism." *Brain and Development* 27 (7):509-16. doi: 10.1016/j.braindev.2005.01.003.

Mirenda, Pat, and Kenneth E. Brown. 2009. "A Picture is Worth a Thousand Words: Using Visual Supports for Augmented Input with Individuals with Autism Spectrum Disorders." In *Autism Spectrum Disorders and AAC*, edited by Pat Mirenda and Teresa Iacono, 303-332. Baltimore, Maryland: Paul H. Brookes.

Miro, Elena, M.ª del Carmen Cano-Lozano, and Gualberto Buela-Casal. 2002. "Electrodermal Activity during Total Sleep Deprivation and its Relationship with other Activation and Performance Measures." *Journal of Sleep Research* 11 (2):105-12. doi: 10.1046/j.1365-2869.2002.00286.x.

Moudry Quilty, Kimberly. 2007. "Teaching Paraprofessionals How to Write and Implement Social Stories for Students With Autism Spectrum Disorders." *Remedial and Special Education* 28 (3):182-189. doi: 10.1177/07419325070280030701.

MPI, (Max Planck Institute) for Psycholinguistics. 2016. "ELAN, Linguistic Annotator." The Language Archive, Max Planck Institute for Psycholinguistics. http://tla.mpi.nl/tools/tla-tools/elan/.

Nagai, Yoko, and Hugo D. Critchley. 2008. "Novel Therapeutic Application of Galvanic Skin Response (GSR) Biofeedback to a Neurological Disorder: Mechanisms underlying Biofeedback in Epilepsy Management." In *Epilepsy Research Progress*, edited by K. Takahashi, 175-205. New York, New York: Nova Science Publishers.

Nagai, Yoko, Hugo D. Critchley, Eric Featherstone, Michael R. Trimble, and Raymond J. Dolan. 2004. "Activity in Ventromedial Prefrontal Cortex Covaries with Sympathetic Skin

Conductance Level: A Physiological account of a "Default Mode" of Brain Function." *Neuroimage* 22 (1):243-51. doi: 10.1016/j.neuroimage.2004.01.019.

Nelson, Heidi D., Peggy Nygren, Miranda Walker, and Rita Panoscha. 2006. "Screening for Speech and Language Delay in Preschool Children: Systematic Evidence Review for the US Preventive Services Task Force." *Pediatrics* 117 (2):e298-e319. doi: 10.1542/peds.2005-1467.

Ollendick, Thomas H., and Jerome A. Cerny. (1981) 2013. *Clinical Behavior Therapy with Children*. Reprint, New York, New York: Springer Science & Business Media.

OMsignal. 2016. "OMsignal bra." Last Accessed March, 2016. http://omsignal.com/pages/omsignal-bra.

Payne, Andrew FH, Michael E Dawson, Anne M Schell, Kulwinder Singh, and Christopher G Courtney. 2013. "Can you give me a hand? A comparison of hands and feet as optimal anatomical sites for skin conductance recording." *Psychophysiology* 50 (11):1065-1069. doi: 10.1111/psyp.12093.

Picard, Rosalind W., Szymon Fedor, and Yadid Ayzenberg. 2016. "Multiple Arousal Theory and Daily-Life Electrodermal Activity Asymmetry." *Emotion Review* 8 (1):62-75. doi: 10.1177/1754073914565517.

Pinborough-Zimmerman, Judith, Robert Satterfield, Judith Miller, Deborah Bilder, Shaheen Hossain, and William McMahon. 2007. "Communication Disorders: Prevalence and Comorbid Intellectual Disability, Autism, and Emotional/Behavioral Disorders." *American Journal of Speech-Language Pathology* 16 (4):359-367. doi: 10.1044/1058-0360(2007/039).

Poh, Ming-Zher , Tobias Loddenkemper, Nicholas C. Swenson, Shubhi Goyal, Joseph R. Madsen, and Rosalind W. Picard. 2010. "Continuous Monitoring of Electrodermal Activity during Epileptic Seizures using a Wearable Sensor." Proceedings of the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS), Buenos Aires, Argentina, August 31 - September 4.

Poh, Ming-Zher, Tobias Loddenkemper, Claus Reinsberger, Nicholas C. Swenson, Shubhi Goyal, Mangwe C. Sabtala, Joseph R. Madsen, and Rosalind W. Picard. 2012. "Convulsive Seizure Detection using a Wrist-worn Electrodermal Activity and Accelerometry Biosensor." *Epilepsia* 53 (5):e93-e97. doi: 10.1111/j.1528-1167.2012.03444.x.

Poh, Ming-Zher, Nicholas C. Swenson, and Rosalind W. Picard. 2010. "A Wearable Sensor for Unobtrusive, Long-Term Assessment of Electrodermal Activity." *IEEE Transactions on Biomedical Engineering* 57 (5):1243-1252. doi: 10.1109/TBME.2009.2038487.

Qi, Cathy Huaqing, and Ann P. Kaiser. 2004. "Problem Behaviors of Low-Income Children With Language Delays: An Observation Study." *Journal of Speech, Language, and Hearing Research* 47 (3):595-609. doi: 10.1044/1092-4388(2004/046).

Ramgopal, Sriram, Sigride Thome-Souza, Michele Jackson, Navah E. Kadish, Iván Sánchez Fernández, Jacquelyn Klehm, William Bosl, Claus Reinsberger, Steven Schachter, and Tobias Loddenkemper. 2014. "Seizure Detection, Seizure Prediction, and Closed-loop Warning Systems in Epilepsy." *Epilepsy & Behavior* 37:291-307. doi: http://dx.doi.org/10.1016/j.yebeh.2014.06.023.

Redmond, Sean M., and Mabel L. Rice. 1998. "The Socioemotional Behaviors of Children with SLI: Social Adaptation or Social Deviance?" *Journal of Speech, Language, and Hearing Research* 41 (3):688-700. doi: 10.1044/jslhr.4103.688.

Reichow, Brian, Erin E. Barton, Leslie Good, and Mark Wolery. 2009. "Brief Report: Effects of Pressure Vest Usage on Engagement and Problem Behaviors of a Young Child with Developmental Delays." *Journal of Autism and Developmental Disorders* 39 (8):1218-1221. doi: 10.1007/s10803-009-0726-3.

Riobo, Ivan, A. Parnami, Javier Hernandez, and Gregory D. Abowd. 2014. "G.L.I.M.: Glass Live Interaction Monitor, Live Internal State Interaction Monitor using Google Glass + EDA." In Extended Abstracts of the International Meeting for Autism Research (IMFAR), Atlanta, Georgia, May 14-17.

Román, Francisco, Francisco A. García-Sánchez, José M. Martínez-Selva, Jesús Gómez-Amor, and Eduvigis Carrillo. 1989. "Sex differences and Bilateral Electrodermal Activity." *The Pavlovian Journal of Biological Science* 24 (4):150-155. doi: 10.1007/bf02964598.

Sano, Akane, and Rosalind W. Picard. 2013. "Stress Recognition Using Wearable Sensors and Mobile Phones." Humaine Association Conference on Affective Computing and Intelligent Interaction (ACII), Geneva, September 2-5.

Sinclair, Jim. 1999. "Why I Dislike "Person First" Language." Autism Mythbusters. http://autismmythbusters.com/general-public/autistic-vs-people-with-autism/jim-sinclair-why-i-dislike-person-first-language/.

Snodgrass, Melinda R., Jennifer G. Kim, Karrie Karahalios, Mindy Borden, and James W. Halle. 2015. "The Effects of a Pressure Vest on Task Engagement, Challenging Behavior, and a Physiological Measure of Stress for a Child with Intellectual Disability." University of Illinois. Illinois Digital Environment for Access to Learning and Scholarship. http://hdl.handle.net/2142/79051.

Snow, Kathie. 2009. "To Ensure Inclusion, Freedom, and Respect for All, it's Time to Embrace People First Language." https://www.disabilityisnatural.com/pfl-articles.html.

Spire. 2015. "Spire." Last Accessed, March 2015. https://www.spire.io/.

Squire, Larry R., Floyd E. Bloom, Nicholas C. Spitzer, Sascha du Lac, Anirvan Ghosh, and Darwin Berg, eds. 2008. *Fundamental Neuroscience*. Third ed. Oxford, United Kingdom: Elsevier.

Stevens, Sheila, and John Gruzelier. 1984. "Electrodermal Activity to Auditory Stimuli in Autistic, Retarded, and Normal Children." *Journal of Autism and Developmental Disorders* 14 (3):245-260. doi: 10.1007/bf02409577.

Stout, A. 2015. "These Wristbands can Detect Meltdowns Before they even Happen." The Autism Site, Last Modified October. http://blog.theautismsite.com/bio-sensor/.

Swaggart, Brenda L., Elisa Gagnon, Stacey Jones Bock, Theresa L. Earles, Colleen Quinn, Brenda Smith Myles, and Richard L. Simpson. 1995. "Using Social Stories to Teach Social and Behavioral Skills to Children with Autism." *Focus on Autism and Other Developmental Disabilities* 10 (1):1-16. doi: 10.1177/108835769501000101.

Tomchek, Scott D., and Winnie Dunn. 2007. "Sensory Processing in Children with and without Autism: A Comparative Study using the Short Sensory Profile." *American Journal of Occupational Therapy* 61 (2):190-200. doi: 10.5014/ajot.61.2.190.

Turpin, Graham, Phil Shine, and Malcolm Lader. 1983. "Ambulatory Electrodermal Monitoring: Effects of Ambient Temperature, General Activity, Electrolyte Media, and Length of Recording." *Psychophysiology* 20 (2):219-224. doi: 10.1111/j.1469-8986.1983.tb03291.x.

Vetrugno, Roberto, Rocco Liguori, Pietro Cortelli, and Pasquale Montagna. 2003. "Sympathetic Skin Response." *Clinical Autonomic Research* 13 (4):256-270. doi: 10.1007/s10286-003-0107-5.

Violeta, Athea. 2015. "Doctor creates wristband that predicts outbursts in individuals with autism." Autism Daily Newscast, Last Modified September 23. http://www.autismdailynewscast.com/doctor-creates-wristband-predicts-outbursts-individuals-autism/31274/altheavioleta/.

Wallin, B. Gunnar. 1981. "Sympathetic Nerve Activity Underlying Electrodermal and Cardiovascular Reactions in Man." *Psychophysiology* 18 (4):470-476. doi: 10.1111/j.1469-8986.1981.tb02483.x.

Wesley, Naissan O., and Howard I. Maibach. 2003. "Racial (Ethnic) Differences in Skin Properties." *American Journal of Clinical Dermatology* 4 (12):843-860. doi: 10.2165/00128071-200304120-00004.

Williams, Leanne M., Kerri J. Brown, Pritha Das, Wolfram Boucsein, Evgeni N. Sokolov, Michael J. Brammer, Gloria Olivieri, Anthony Peduto, and Evian Gordon. 2004. "The Dynamics of Cortico-amygdala and Autonomic Activity over the Experimental Time Course of Fear Perception." *Cognitive Brain Research* 21 (1):114-123. http://dx.doi.org/10.1016/j.cogbrainres.2004.06.005.

Yerkes, Robert M., and John D. Dodson. 1908. "The Relation of Strength of Stimulus to Rapidity of Habit-formation." *Journal of Comparative Neurology and Psychology* 18 (5):459-482. doi: 10.1002/cne.920180503.

CHAPTER 4:

Nonshared Environmental Influences on Language Development: A Monozygotic Twin Differences Study

Abstract

Despite support for a range of genetic and environmental influences on child language development, studies within the field of communication sciences and disorders remain largely focused on the study of maternal linguistic input. Results from such studies are most often correlational in nature and do not control for the myriad of confounding factors including potential genetic effects. In contrast, behavioral genetic studies that control for systematic genetic differences have revealed significant nonshared environmental (NSE) influences on language development, but have not identified what those specific factors might be. The present study builds on prior behavioral genetic studies of language (DeThorne et al., 2008; Harlaar, DeThorne, Mahurin-Smith, Aparicio Betancourt, & Petrill, 2016) and offers to the best of our knowledge, the first application of the monozygotic (MZ) twin differences method to understanding nonshared environmental or "person-specific" influences on children's language development. This work utilizes a rich longitudinal twin database, the Western Reserve Reading and Math Project (WRRMP; Petrill, Deater-Deckard, Thompson, DeThorne, & Schatschneider, 2006), to investigate the extent to which discordance in four twin-specific environmental measures (i.e., birthweight, breastfeeding, and home reading exposure and parenting at mean age 7) are related to differential language outcomes in monozygotic twins at two time points: mean ages 10 (n = 115 pairs) and 12 years (n = 108 pairs). Language outcomes focus on both productive measures from narrative language measures (i.e, Productive language) and formal measures from standardized language measures (i.e., Formal language). Across time points and forms of language assessment, at least one significant association emerged between language discordance and discordance in a) birthweight, b) self-reported parenting, and c) home reading exposure, with stronger and more consistent effects for Formal language. For the full unselected sample, only birthweight discordance at mean age 12 replicated its effect across both Formal and Productive language, and only parent self-reported differential negativity replicated its effect across both time points (for Formal language only). NSE effects were also moderated by the extent of MZ language discordance, with effects increasing as a function of increased language discordance. Potential mechanisms are discussed within the context of the complex and dynamic interplay between genetic and environmental influences over time. Implications for future research are also discussed.

4.1 Introduction

Language development is key for social and academic success; in fact, promoting language growth has been characterized as a national priority (Jacobson, 2001). Early language delays or impairments have been associated with negative long-term consequences. For example, children with language impairments are more likely to have socioemotional and behavioral difficulties, and language impairment has been negatively associated with literacy acquisition (Forget-Dubois et al., 2009; McCardle, Scarborough, & Catts, 2001; Qi & Kaiser, 2004; Redmond & Rice, 1998). Given the role of language in communication and in other developmental domains, it becomes important to understand influential factors on child language development.

Aside from genetics, individual differences have been largely associated with nonshared environmental influences within behavioral genetic designs (Plomin, Asbury, & Dunn, 2001). However, the study of candidate nonshared environmental influences on language development remains underdeveloped, particularly within the literature in communication sciences and disorders (CSD). In a recent review of the CSD literature (Rogers, Nulty, Aparicio Betancourt, & DeThorne, 2015) we examined what causal influences on child language development were explicitly being studied by reviewing 2,921 abstracts published between 2003-2013 across five journals. Of the eligible articles (n = 346), the majority addressed environmental influences (83%), with the remaining articles addressing genetic influences either in isolation (11%) or in concert with environmental influences (6%). Of those that addressed environmental influences, the majority focused on therapist intervention (52%), caregiver linguistic input (22%), and caregiver qualities such as socioeconomic status (13%). Only 8% of the abstracts addressed other environmental influences such as diet and preterm birth (e.g., Zubrick, Taylor, Rice, & Slegers, 2007). A more in-depth review of all the eligible articles published in 2013 (n = 34) showed most studies controlled for variables related to caregiver qualities, and only 9% (3/34) of the studies addressed the complex interplay between genetic and environmental influences. This literature review highlighted the need to study a broader range of environmental influences on child language development, and to consider the complex interplay between genetic and environmental influences.

To address this pressing gap in the CSD literature, the present study will use a longitudinal, genetically-sensitive twin design to examine candidate nonshared environmental (NSE) influences that have been previously associated with developmental outcomes. For the present study, we

associate NSE with any variable that can generate difference scores within a monozygotic (MZ) twin pair. Consequently, factors such as parenting, which are shared across twins still have the potential to be differentially constructed across individual twins, thereby potentially contributing to individual nonshared effects. Specifically, we will examine the influence of birthweight [16], extent of breastfeeding, home reading exposure, and parenting on school-age language development as measured by standardized testing and narrative language samples.

4.1.1 Twin methodology. Genetically sensitive designs, specifically use of MZ twin difference analyses, provide a unique opportunity to identify NSE effects by controlling for both genetic and shared environmental influences in order to examine child-specific variation (Asbury, Dunn, & Plomin, 2006a; Mullineaux, Deater-Deckard, Petrill, & Thompson, 2009; Pike, Reiss, Hetherington, & Plomin, 1996). MZ twins share approximately 100% of their DNA in addition to having indistinguishable epigenomes early in development (Fraga et al., 2005), and sharing many perinatal (e.g., the majority share the placenta) and postnatal factors. On the other hand, dizygotic twins (DZ) share approximately 50% of their DNA in addition to similar perinatal and postnatal factors, with the majority having separate placentae. Twin similarity increases as a function of common genetic and shared environmental factors, and decreases as a function of nonshared environmental experiences. Approximately 1/3 of twins are monozygotic and 2/3 are dizygotic. Specifically, DZ twinning is associated with genetic factors, increased maternal age, race/ethnicity, fertility treatments, among other factors, whereas the etiology of MZ twins is less clear (Stromswold, 2001, 2006).

Consistent with the high heritability of language development, MZ twins have higher concordance of speech and language disorders (e.g. autism, specific language impairment, speech sound disorders) than DZ twins (Lewis & Thompson, 1992). If speech-language development were solely influenced by genetic and shared environmental factors, we would expect 100% concordance. However, there is still a high degree of discordance among MZ twin pairs which is driven, in part, by external nonshared environmental factors (Wong et al., 2014). Specific to language, twin heritability estimates are usually below 60% and some MZ twins can have substantially different linguistic profiles (Stromswold, 2006). An MZ twin differences method,

[16] Birthweight, although often perceived as an outcome variable, is being used as a proxy for potential NSE influences.

which controls for both genetic effects and shared environmental effects, can be used to elucidate candidate nonshared environmental factors that may in part drive MZ twin linguistic discordance.

More broadly, genetic and environmental factors can affect linguistic development and cause MZ twins to be linguistically discordant. Although the majority of MZ twins are genetically identical, a minority of MZ twins have different genotypes due to factors such as chromosomal non-disjunction that may lead to one twin having Down syndrome and the other remaining unaffected, and spontaneous mutations. In regards to environmental influences, MZ twins prenatal environment can vary depending on when the zygote divides into two identical zygotes. The majority of MZ twins share a placenta (i.e., monochorionic) and have separate amniotic sacs (70-75%), with 20-25% having separate placentae and amniotic sacs, and a minority sharing placenta and amniotic sacs (1-5%). Such differences may have genetic and perinatal environmental implications. For example, those with separate placentae and amniotic sacs usually split earlier in development and have a greater likelihood of different spontaneous mutations (Stromswold, 2006).

Finally, epigenetic differences are associated with MZ differences and have been reported to increase with age as a result of both internal (e.g., small differences in transmitting epigenetic information), and external environmental factors (e.g., nutrition, physical activity). Epigenetic processes refer to potentially heritable modifications of gene expression without altering the genome, via for example, methylation, gene silencing, or x chromosome inactivation by micro-RNA regulation. To illustrate, x chromosome inactivation patterns have been shown to be more similar in monochorionic MZ twins, who share a more similar perinatal environment, compared to dichorionic MZ twins (Stromswold, 2006). MZ twins with increased nonshared environments have increased differences in their epigenomes and more phenotypic differences, compared to those who have spent more of their lives together and with more similar lifestyles (Fraga et al., 2005). Although MZ twins originally share the same genotype and have indistinguishable epigenomes, their environments become more and more different as they age. By adulthood, MZ twins' epigenome is very different; given the difference in gene expression patterns, they are no longer technically identical (Fraga et al., 2005; Poulsen, Esteller, Vaag, & Fraga, 2007; Wong, Gottesman, & Petronis, 2005). Nonshared environmental factors such as diet, adverse health outcomes, parental interactions, exposure to toxins, and stress may in part drive these epigenetic differences and ultimately influence their language development. The MZ differences approach provides a simple and sensitive method that can be used to examine such factors.

Based on a review of twin studies conducted by Plomin and Kovas (2005), estimates of genetic effects on child language abilities have varied widely, ranging from 16% to 100% (see also Stromwold, 2001). The variability in estimates has been attributed to a broad range of factors including age (e.g., DeThorne, Harlaar, Petrill, & Deater-Deckard, 2012; Harlaar, DeThorne, Mahurin-Smith, Aparicio Betancourt, & Petrill, 2016), language domain (e.g., Stromswold, 2001), degree of linguistic ability (e.g., Stromswold, 2001, 2006), and form of assessment (e.g., DeThorne et al., 2008; Harlaar et al., 2016). We emphasize here two prior WRRMP studies focused on school-age children and multiple forms of language assessment that together form the groundwork for the present work. DeThorne and colleagues (2008) conducted a multivariate genetic analysis in 380 seven-year-old twins during the second wave of the WRRMP (i.e., Home visit 2). The multiple language measures loaded on a Conversational latent factor and a Formal latent factor. The Conversational factor included measures taken from conversational language samples, including mean length of utterance (MLU), number of total words (NTW), and number of different words (NDW). The Formal factor included two standardized vocabulary assessments: the Boston Naming Test, and the Vocabulary subtest from the Stanford-Binet Intelligence Scale. Multivariate analyses revealed a heritability of .70 for the Conversational factor and .45 for the Formal factor; nonshared environmental effects were negligible for the Formal factor and .30 for the Conversational factor, albeit nonsignificant.

In a follow-up study from WRRMP, Harlaar et al., (2016) used structural equation modeling to examine the longitudinal genetic and environmental contributions to individual differences in children's language skills (n = 498 to 522) at mean ages 10, 11, and 12 (home visits 5, 6 and 7 respectively). Similar to DeThorne and colleagues (2008), measures loaded on a Productive and a Formal latent factor. The Productive factor was based on narrative language measures and included MLU, NTW, and NDW, and the Formal factor included standardized-test scores from the Test of Narrative Language and three subtests of the Clinical Evaluation of Language Fundamentals – Fourth Edition. Specifically, the predominate influence on the Productive factor was nonshared effects across time points (55-90%), with no significant genetic or shared environmental effects. The Formal factor demonstrated strong significant genetic effects across time points (82-86%), with small but significant nonshared effects across time points (5-6%). The Productive factor showed limited but significant longitudinal stability in NSE effects across time points, and the Formal factor showed high longitudinal stability in the genetic and NSE effects across time points.

This increase in NSE on language at early adolescence relative to earlier ages (i.e., DeThorne et al., 2008, 2012) further supported the need to identify potential nonshared environmental influences. The remainder of this section will focus on candidate nonshared environmental factors on language development organized as peri/early postnatal and social/linguistic.

4.1.2 Perinatal and early postnatal factors. The pre- , peri-, and post-natal periods are sensitive periods of both opportunity and vulnerability, with brain development beginning in the 3rd week of gestation (Stiles, 2008). As such, factors such as diet, exposure to teratogens (e.g. cigarette smoking), and infection, may significantly shape neurodevelopment through influencing the blood-brain barrier, synaptogenesis, proliferation of oligodendrocytes and myelination, and pruning. Cortical networks begin to form prenatally and continue to do so postnatally, with a large growth spurt of neural networks, including those involved in cognition, forming during the first few months of life (Stiles, 2008). Neural plasticity is particularly evident in the immature cerebral cortex, and is most marked in regions associated with higher cortical functions such as language (Huttenlocher, 2002). Spatially specific molecular signaling of neural progenitor cells, together with activity-dependent signaling will shape the final organization and functions of the neurons within the neocortex. The formation of synapses between neurons (i.e., synaptogenesis), for example, is shaped by post-natal sensory and motoric experiences after birth, particularly during the first three years of life, and declines thereafter with the onset of pruning. Through pruning, which involves the removal of extra neurons and synaptic connections via apoptosis, neuronal networks enabling specific functions emerge (Stiles, 2008). The development of the respective neural networks, which are thought to underlie information processing in the cerebral cortex, is determined to a significant extent by environmental influences, especially during the postnatal period (Huttenlocher, 2002). To illustrate, although children continue to be able to acquire language later in life, children's ability to acquire language declines after the decline of synaptogenesis, likely due to decreased brain plasticity.

The present study examines the potential influence of a perinatal (i.e. birthweight) and an early postnatal factor (i.e., extent of breastfeeding) in later language development. These factors may potentially influence language development via direct alterations to brain anatomy and physiology and indirect influences on language learning.

4.1.2.1 Birthweight.

4.1.2.1.1 Background. Low and high birthweight, as well as moderate to severe birthweight discordance in twin pregnancies are associated with higher rates of perinatal morbidity and mortality (Ananth, Demissie, & Hanley, 2003; Boulet, Alexander, Salihu, & Pass, 2003; Fanaroff et al., 2007; Wen et al., 2005), as well as long-term complications, chronic diseases, and developmental delays (United Nations Children's Fund [UNICEF] & World Health Organization [WHO], 2004). Low birthweight (LBW) is defined as weight at birth of less than 2.5 kg, with very low birthweight (VLBW) defined as weight at birth of less than 1.5 kg (Fanaroff et al., 2007; Martin, Hamilton, Osterman, Driscoll, & Mathews, 2017; UNICEF & WHO, 2004). High birthweight or fetal macrosomia is defined as greater than 4.0 kg, with increased health risks when greater than 4.5 kg (Abel et al., 2013; Boulet et al., 2003; Mayo Clinic, 2015). Finally, moderate to severe discordant growth in twins is defined as greater than 15%, with increased health risks when ≥ 20% (Ananth et al., 2003; Armson et al., 2006; Hartley & Hitti, 2005; Wen et al., 2005). Approximately 9% of infants worldwide are diagnosed with fetal macrosomia (Mayo Clinic, 2015), and approximately 15.5% of all newborns worldwide are born with low birthweight (UNICEF & WHO, 2004; see also Martin et al., 2015). Over 50% of multiple birth infants have low birthweight compared to 6% of singletons (Martin et al., 2015). The lower or higher the birthweight the greater the risk for complications.

The primary cause of low birthweight is premature birth (born before 37 weeks of gestation), with the second leading cause being fetal intrauterine growth restriction (IUGR), or slower than typical velocity of growth (UNICEF & WHO, 2004). Although birthweight is closely associated with premature birth, it has been reported to be an independent predictor of both short-term and long-term outcomes, including linguistic delays and impairments (Fanaroff et al., 2007; Malin, Morris, Riley, Teune, & Khan, 2014; Matthews, MacDorman, & Thoma, 2015; Stromswold, 2006). Low birthweight is likely a result of a myriad of factors that relate to the infant, the mother, and the physical environment (UNICEF & WHO, 2004). Factors associated with the infant and the mother include multibirth pregnancies (e.g. twins weigh less than singletons), and maternal and fetal genetic make-up (e.g. trisomy 18) (Varner & Esplin, 2005). The mother's body composition at conception also plays a role and is known to be influenced by her own fetal growth, her diet and health from birth to pregnancy, maternal age (with increased risk in women younger than 20 years and older than 35; Fraser, Brockert, & Ward, 1995; U.S. Department of Health and

Human Services [HHS], 2014), low pre-pregnancy weight, and short interval between pregnancies. Finally, the mother's lifestyle, exposures, and complications during pregnancy are also associated with LBW. Examples of potential factors include alcohol consumption, maternal smoking, prolonged high-altitude exposure, low maternal weight gain, physical work, exposure to malaria, hypertension, pregnancy healthcare, and maternal or fetal stress (UNICEF & WHO, 2004).

Specific to discordant growth in twin pregnancies, a higher incidence of fetal distress, requirement for supplemental oxygen, respiratory distress syndrome, and low Apgar scores are associated with increased growth discordance (Hartley & Hitti, 2005). In addition to the low birthweight etiologies, factors impacting discordant growth in twins are often associated with intrauterine growth restriction, structural and functional placental abnormalities, and twin-twin transfusion syndrome (TTTS) (Siddiqui & McEwan, 2007). Approximately 15 - 20% of MZ twins suffer from TTTS, in which one twin donates fetal blood to the other twin. Both decreased blood flow in the donor twin, and increased blood flow in the recipient twin are associated with adverse health outcomes including increased risk of brain injuries (Stromswold, 2006).

Contrary to the low birthweight literature, the relationship between macrosomia and later language outcomes remains relatively unexplored. Some studies report associations between gestational diabetes, a risk factor for fetal macrosomia, and poorer language outcomes (e.g., Dionne, Boivin, Séguin, Pérusse, & Tremblay, 2008; Perna, Loughan, Le, & Tyson, 2015), and one study reported an increased risk of autism spectrum disorders in macrosomic infants, independent of prematurity (Abel et al., 2013). Given limited research, the remainder of this section will focus on the potential influence of low birthweight on linguistic skills. Given the prevalence of low birthweight worldwide, it is particularly important to examine the long-term neurodevelopmental sequelae.

4.1.2.1.2 Potential mechanisms. The question remains as to how low birthweight may lead to subsequent language difficulties. Potential mechanisms can be broadly divided into the two general categories, which may not be mutually exclusive. First, direct insults or alteration to fetal brain anatomy and physiology may lead to overt or covert damage to the neural networks involved with language. These include: a) reduced brain volume as a result of low birthweight and/or premature birth leading to decreased general cognitive functioning (e.g., de Kieviet, Zoetebier, Van Elburg,

Vermeulen, & Oosterlaan, 2012), b) overt or covert hypoxic/ischemic[17] brain injuries such as intraventricular hemorrhage and periventricular white matter lesions (e.g., see Whitaker et al., 1996), which may occur as a result of decreased oxygen-carrying capacity (e.g., blood clotting disorders), and/or a dysfunctional oxygen-delivery system which may occur in cases of e.g., diabetes, or placental or cord complications [18] (HHS, 2014; Severi et al., 2000; Stromswold, 2006) and finally c), brain physiology may be influenced by toxic exposure[19], malnutrition, and infections such as cytomegalovirus (Stromswold, 2006; UNICEF & WHO, 2004). Such direct influences on brain development may lead to the disruption of fibers that mediate higher cortical functions such as language. In fact, perisylvian language areas are particularly vulnerable to hypoxic/ischemic brain injuries given they are located in a vascular watershed and therefore are at a greater risk for hypoperfusion (Stromswold, 2006).

The second potential mechanism focuses on indirect influences on language learning via postnatal complications and psychosocial factors. Activity or experience-dependent signaling is known to shape neural networks (Huttenlocher, 2002; Stiles, 2008). Differences in postnatal experiences relevant to language development during a sensitive period of brain maturation could therefore influence linguistic ability. Infants with low birthweight are at greater risk for adverse perinatal and postnatal health outcomes (Malin et al., 2014; Ramachandrappa & Jain, 2009)[20]. Such adverse health outcomes often lead to increased hospital stay and increased parental stress, which are associated with differences in caregiver-infant interactions (Korja, Latva & Lehtonen, 2012; Ramachandrappa & Jain, 2009). Children's health may influence parents' feelings towards their children and may represent a source of differential parental treatment (e.g., Caspi et al., 2004). Prolonged hospitalization may also lead to decreased skin-to-skin contact; skin-to-skin contact has been found to be beneficial to infants, including low birthweight infants (Whitelaw, Heisterkamp,

[17] Hypoxia refers to reductions of oxygen; ischemia refers to diminished blood supply.
[18] Decreased oxygen-carrying capacity is associated with conditions such as cyanotic heart disease, thrombophilias or other blood clotting disorders, smoking, and hemoglobinopathy. Dysfunctional oxygen-delivery system may occur in cases of diabetes, maternal hypertension, and some autoimmune diseases. Placental or cord complications could also lead to maternal bleeding and decreased blood supply to the fetus, decreasing fetal oxygenation and increasing the risk of a hypoxic/ischemic brain injury (e.g. placenta previa, chronic abruption, and abnormal cord insertion) (HHS, 2014; Severi et al., 2000; Stromswold, 2006).
[19] Toxic exposure may be due to medication use, alcohol or other substance misuse, smoking, or hyperbilirubinemia (i.e, jaundice) which is also associated with low birthweight.
[20] E.g., low oxygen levels at birth, inability to regulate body temperature, breathing problems (e.g., RDS), neurological problems (e.g., intraventricular hemorrhage), GI problems (e.g., necrotizing enterocolitis), and SIDS (Sudden Infant Death Syndrome).

Sleath, Acolet, & Richards, 1988). Socialization is key to language development and differences in social interactions in infancy may set the stage for differences in later linguistic development. In addition, exposure to intense sensory stimulation in the neonatal intensive care unit (NICU) has been associated with negative outcomes. The American Academy of Pediatrics Committee on Environmental Health (AAP, 1997) concluded exposure to noise and other environmental factors in the NICU may lead to cochlear damage, and poor growth and development. Another indirect factor that may influence language learning is the increased risk of postnatal complications such as feeding and swallowing difficulties (Bu'Lock, Woolridge, & Baum, 1990; Ramachandrappa & Jain, 2009). Such difficulties may lead to improper nutrition, increased risk of infections, and changes in brain physiology. The increased risk of adverse health outcomes associated with low birthweight may also influence infant's overall learning and memory. Infants who are at clinically significant risk for morbidity and mortality may devote fewer resources for the development of higher-order cognitive processes, including language.

Whether as a result of one or all of the potential mechanisms, preterm birth is associated with less developed mid-temporal and parieto-occipital cortices, which play important roles for auditory, language and other cognitive processes, and for integration of stimuli and information. Mid-temporal cortical volumes have been positively associated with full scale, performance, and verbal IQs (Peterson et al., 2000). Other studies confirm preterm births, very low birthweight infants, and small for gestational age (GA) births (birthweight $<10^{th}$ percentile), are associated with smaller total and regional brain volumes (e.g., hippocampus, corpus callosum), smaller gray/white matter volume, and/or white matter integrity, and decreased cognitive functioning (de Kieviet et al., 2012; Eikenes, Løhaugen, Brubakk, Skranes, & Håberg, 2011; Martinussen et al., 2009). Neurological differences may in turn influence how infants interact with their environment and how others interact with them, which in turn shapes brain development.

4.1.2.1.3 Empirical research. Low birthweight as measured by medical records or parental recall has been correlated with subsequent long-term difficulties in speech, language, attention, general cognitive skills, social skills, behavior difficulties, and with disorders such as attention-deficit/hyperactivity disorder, autism spectrum disorders, and learning disabilities (Abel et al., 2013; Aylward, 2002; Grunau, Whitfield, & Davis, 2002; Jansson-Verkasalo et al., 2004; Johnson & Breslau, 2000; Mahurin-Smith, DeThorne, Logan, Channell, & Petrill, 2014; Pharoah, Stevenson, Cooke, & Stevenson, 1994; Veen et al., 1991). Lung problems and vision and hearing

loss have also been reported. Of interest, fetal growth restriction is also associated with lower IQ in the general population (Eriksen, Sundet, & Tambs, 2010). Specific to language development, LBW infants are reported to have elevated risks for speech and language impairments, at least until school age (Grunau, Kearney, & Whitfield, 1990; Jennische & Sedin, 2001). Whether children without frank neurological impairment continue to present with language difficulties at school age, is less clear. To illustrate, I highlight two representative studies, the first focuses on preschool language skills and the second on school-age language.

Jansson-Verkasalo and colleagues (2004) conducted a group study examining language in 17 VLBW preterm children (BW < 1,500g, GA < 34 weeks) and 17 matched controls (Mean BW = 3,617 g, Mean GA = 39.7 w) at age 2 and at age 4. Birthweight data were collected via medical records. The research design benefitted from the use of a longitudinal design that assessed language through both standardized measures and language samples. VLBW children showed lower receptive language scores and used shorter and more immature productions than matched controls at age 2 as measured by the Reynell Developmental Language Scales, the MacArthur-Bates Communicative Development Inventory, and 15-min video-recorded language samples during free play between the child and the caregiver. By four years of age, VLBW children showed significantly decreased language comprehension, vowel and consonant discrimination, and word production compared to their matched controls, as measured by the Reynell Developmental Language Scales and a battery of different tests over similar domains. Performance in all the language assessments for the VLBW children at age 2 predicted language performance at age 4; correlations were weaker or nonexistent for the control group (readers are referred to similar findings in Grunau et al., 1990; Jennische & Sedin, 2001).

Similar to Jansson-Verkasalo and colleagues (2004), Mahurin-Smith and colleagues (2014) examined the impact of parent reported prematurity (<= 32 weeks) or VLBW (< 1,500 g) on language skills at school age (ages 7, 8 and 10 years) using standardized and language sample measures within the context of the Western Reserve Reading and Math Project (WRRMP). The premature or VLBW group (n = 57) was generally outperformed by full-term matched controls (n = 57), albeit only language outcomes as measured by standardized assessments were significantly different. No statistically significant differences were seen in language skills across groups for the language sample measures. The authors note that none of the study participants had extremely low birthweight or overt neurological impairment.

The long-term developmental trajectory of children with low birthweight is highly heterogeneous (Stromswold, 2006). The heterogeneity in cognitive outcomes, such as language, associated with a child's birthweight may be due to the severity of perinatal complications (e.g., leading to intellectual disabilities), predisposition to cognitive impairment, and to moderating factors that may mitigate the impact of birthweight on the developing brain. Previously proposed moderating factors include parent-child interactions (to be discussed in parenting section), diet (e.g. a diet rich in long-chain fatty acids, to be discussed in breastfeeding section), accessibility to therapy, parent education, exposure to literacy-rich activities (to be discussed in reading section), and exposure to toxins among other environmental variables.

4.1.2.2 Breastfeeding.

4.1.2.2.1 Background. The American Academy of Pediatrics (2012) and the World Health Organization (WHO, 2001) recommend exclusively breastfeeding infants for the first 6 months of life and continuing to breastfeed up to 1 or 2 or more years of age with adequate complementary foods, while also acknowledge the need to consider mother-infant preferences. Breastfeeding is associated with a plethora of short and long-term health benefits for both the mother and the infant. For example, infants who are breastfed have a decreased risk of obesity (Burke et al., 2005; Harder, Bergmann, Kallischnigg, & Plagemann, 2005), diabetes (Sadauskaitė-Kuehne, Ludvigsson, Padaiga, Jašinskienė, & Samuelsson, 2004), infections (Galton Bachrach, Schwarz, & Bachrach, 2003; Duncan et al., 1993; Mårild, Hansson, Jodal, Oden, & Svedberg, 2004; Silfverdal, Bodin, & Olcén, 1999), and a higher rate of survival during their first year of life (Bahl et al., 2005; WHO, 2009). Some studies also suggest it may be advantageous for an infant's mental health (Oddy et al., 2010), cognitive development (e.g., Anderson, Johnstone, & Remley, 1999; Kramer et al., 2008; Lucas, Morley, & Cole, 1998; Nyaradi, Oddy, Hickling, Li, & Foster, 2015), and later language development (Leventakou et al., 2015; Oddy et al, 2011; Quigley et al., 2012; for a review see Mahurin-Smith, 2015).

4.1.2.2.2 Potential mechanisms. The influence of an infant's diet, or breastfeeding more specifically, could potentially influence language development via direct alterations to brain anatomy and physiology and indirect influences on language learning, none of which are mutually exclusive. First, breast milk may offer key nutritional building blocks for brain development that are not fully emulated in infant formula, thereby directly influencing brain structure and function

and modulating gene expression. The neurodevelopmental advantage associated with breast milk is thought to be linked primordially to the milk's long-chain polyunsaturated fatty acids (LCPUFAs), particularly docosahexaenoic acid (DHA) (Innis, 2007; McFadyen, Farquharson, & Cockburn, 2003), and to a lesser extent arachidonic acid or ARA (WHO, 2009). DHA is the most abundant fatty acid in the brain, and DHA levels in the brain have been associated with an individual's diet. It plays a role in neurogenesis, neurotransmission, gene expression regulation and oxidative stress inhibition. Decreased DHA in brain tissue has been associated with poorer performance in learning and memory tasks in rodents (Innis, 2007). Human studies also support the hypothesis that dietary fatty acid content influences the structure and function of neural tissue (Deoni et al., 2013; Kafouri et al., 2012). Early nutritional influences in brain structure and function together with early nutritional influences in gene expression (e.g., Farquharson, Jamieson, Logan, Cockburn, & Patrick, 1992; Salem et al., 2001) could have long-term implications to language development (e.g., Anderson et al., 1999), particularly in those infants otherwise predisposed to have language difficulties (e.g., Schultz et al., 2006; see also Lucas, Morley, Cole, Lister, & Leeson-Payne, 1992 and Quigley et al., 2012 for stronger effects of breastfeeding in at-risk populations more broadly). Together with LCPUFAs, some or all of the ingredients in human breast milk may work synergistically to provide neurodevelopmental advantages. In addition to fats, breast-milk also contains water, carbohydrates, protein, vitamins, minerals, anti-infective factors (e.g., immunoglobulin), and other bioactive factors (e.g., bile-salt stimulated lipase and epidermal growth factor). In contrast with formula, breast milk has lower protein and energy content, and higher content of LCPUFAs, cholesterol, and nondigestible carbohydrate (WHO, 2009).

Second, breastfeeding may influence language development indirectly. Lactation may facilitate mother-child bonding in a way that supports later language development. In addition to providing dedicated time together and skin-to-skin contact, lactation is associated with hormonal changes including increases in maternal oxytocin, the milk-ejection hormone and a prosocial neuropeptide. Oxytocin has been associated with feelings of relaxation (Uvnäs-Moberg, 1996; WHO, 2009), improved learning and memory (Hurlemann et al., 2010; Savaskan, Ehrhardt, Schulz, Walter, & Schächinger, 2008; Tomizawa et al., 2003), and even improved emotion recognition (Guastella et al., 2010), and prosocial behaviors (Meyer-Lindenberg, Domes, Kirsch, & Heinrichs, 2011). Whereas lactation increases maternal oxytocin, mother-infant social

interaction can increase oxytocin levels in the mother and the infant (e.g., Feldman, Weller, Zagoory-Sharon, & Levine, 2007). More specifically, improved maternal sensitivity has been reported in lactating mothers (Britton, Britton, & Gronwaldt, 2006; Kim et al., 2011), and increased maternal sensitivity during infancy is correlated with later language development (Baumwell, Tamis-LeMonda, & Bornstein, 1997) (see parenting section for additional details) and other positive neurocognitive outcomes (Rahkonen et al., 2014).

Furthermore, the antimicrobial, anti-inflammatory properties, and bioactive factors in breast milk play an important role in both short- and long-term immune function and regulation of immunoinflammation, indirectly influencing learning and memory. Breast milk protects the infant's immune system and supports the ontogeny of the infant's own immune system (Field, 2005; Hanson, Korotkova, & Telemo, 2003; Lepage & Van de Perre, 2012). An enhanced immune system will lead to decreased risk of infections, and may allow the infant to devote more resources to learning and brain maturation and fewer resources to fight pathogens and modulate immuno-inflammation. Additionally, infant formula is not sterile and has been found to be contaminated with pathogenic bacteria (e.g., Enterobacter sakazakii) (Forsythe, 2005). The increased likelihood of contamination in infants who are formula-fed, together with the improved immune status and immunoregulation in breastfed infants may act as a measure of protection and indirectly influence learning and memory.

4.1.2.2.3 Empirical research. Specific to language, most studies demonstrate positive, modest but statistically significant differences supporting neurodevelopmental advantages for breastfed vs. formula-fed infants in both typical and atypical populations (e.g., Harrison & McLeod, 2010; Peyre et al., 2014). Although attenuated, significant effects are observed in most studies even after controlling for SES, maternal IQ, maternal education, and maternal vocabulary [21] (e.g., Anderson et al., 1999; Daniels & Adair, 2005; cf. Der, Batty, & Deary, 2006). Moreover, a dose-response effect in which prolonged and exclusive breastfeeding are associated with increased benefits has been reported (e.g., Burke et al., 2005; Daniels & Adair, 2005; Harder et al., 2005; Isaacs et al., 2010; Leventakou et al., 2015; Mortensen, Michaelsen, Sanders, & Reinisch, 2002); with boys being particularly responsive (Nyaradi et al., 2015; Oddy et al., 2011). In typical populations, IQ differences as large as 7.5 points in the Wechsler Abbreviated Scales of Intelligence verbal IQ

[21] Controlling for variables such as maternal vocabulary may partial out some of the positive effects of breastfeeding given its effects on maternal learning and memory.

subscale, consisting of the vocabulary and similarities subtests, at 6.5 years have been reported (Kramer et al., 2008; see Lucas et al., 1992 for an 8 pt. IQ difference in preterm infants). Kramer and colleagues (2008) findings are particularly relevant given the methodological rigor of the study and assessment of breastfeeding influences as old as school-age. Specifically, the authors conducted a follow-up on 13,889 healthy breastfed infants and their mothers using a cluster-randomized trial. Breastfeeding exclusivity and duration was promoted in specific maternity hospitals, whereas the control maternity hospitals continued their previously established practices. The children who received decreased breastfeeding duration and exclusivity scored lower on verbal IQ measures and on teacher evaluations of reading and writing skills[22] (Kramer et al., 2008).

The protective effects of breastfeeding have also been reported in association with younger children (Peyre et al., 2014) and clinical or at-risk populations including children with autism (Al-Farsi et al., 2012; Dodds et al., 2011; Schultz et al., 2006), speech impairment (e.g., Mahurin-Smith & Ambrose, 2013), language impairment (Dee, Li, Lee, & Grummer-Strawn, 2007; Harrison & McLeod, 2010; Tomblin, Smith, & Zhang, 1997), and a history of prematurity or very low birthweight (Belfort et al., 2016). Even though there is support for the positive effects of breastfeeding in at-risk populations, perinatal complications (e.g., low birthweight, low Apgar score) have been found to have a small, albeit significant association with decreased incidence of breastfeeding, in which a higher percentage of mothers do not breastfeed their infants (Tamminen, Verronen, Saarikoski, Göransson, & Tuomiranta, 1983; see also Quigley et al., 2012 and Scott, Binns, Oddy, & Graham, 2006). As such, although breastfeeding is likely to have a positive effect on all infants, at-risk infants are less likely to benefit from such effects given the decreased breastfeeding rates.

In regards to the effects of breastfeeding in low birthweight and preterm infants, Belfort and colleagues (2016) examined associations between breastfeeding duration and exclusivity in very preterm infants and very low birthweight infants who are at increased risk of neurodevelopmental impairments and often need fortified feedings and begin feedings via nasogastric tube, thereby changing the quality of breastfeeding interaction. In addition to the population studied, this study is highlighted given its use of medical records to quantify breast milk intake and assessment of brain development and neurocognitive outcomes during the school-age years (n =180). Greater exposure to breastfeeding during the first 28 days was associated with significantly higher verbal

[22] The teachers were blind to the independent variable.

IQ using the Wechsler Abbreviated Scales of Intelligence at 7 years of age (similar to Kramer et al., 2008), as well as significantly higher scores in other neurocognitive outcomes including full scale IQ, performance IQ, math computations, working memory and motor function. Word reading (WRAT4) and language scores (CELF-IV) also increased with increasing breastfeeding exposure, albeit nonsignificantly. Assessment of breastfeeding duration and daily volume during only the first 28 days remained a notable limitation of this study. Ideally, breastfeeding should have been assessed for at least the first three months. It is possible that the nonsignificance of the language outcome as measured by the CELF-IV was due to the lack of breastfeeding assessment beyond the first 28 days or an effect size that was too small to reach statistical significance.

Although null findings between breastfeeding and neurocognitive outcomes have also been reported in other studies (Colen & Ramey, 2014; Der et al., 2006; see Thorpe, Rutter, & Greenwood, 2003 for null findings in twins), such studies have not included strong measures of breastfeeding exclusivity and/or extent of breastfeeding (see Kramer et al., 2008 and Mahurin-Smith, 2015 for a critical review of such studies). In fact, stronger associations are present with increased breastfeeding exclusivity and increased duration of breastfeeding (Leventakou et al., 2015), with a proposed ≥ 3-month threshold to benefit from the effects of breastfeeding (Dee et al., 2007), although a ≥ 9 month threshold has also been reported in infants at-risk for communication impairments (Harrison & McLeod, 2010; Tomblin et al., 1997). It is also important to consider the a priori differences between mothers who tend to breastfeed and those who do not in regard to potential financial (e.g., socioeconomic status, health care access), physical, social (e.g., education, lactation support), and psychological differences (e.g., see Der et al., 2006).

4.1.3 Social and linguistic factors. Although the degree of neural plasticity decreases with age, the human potential for change is evident throughout life; as such, environmental factors at any point in time can influence behavioral phenotypes. Here we focus on the potential influence of parenting and home reading exposure during childhood on later language development, given their predominance in the CSD literature.

4.1.3.1 Parenting.

4.1.3.1.1 Background. Much of the literature on the effects of parenting focuses on parenting style, which refers to the emotional climate surrounding the interactions between parents and their children, including socialization priorities (e.g., values, attitudes) and specific parenting practices

(e.g., behaviors, strategies). Parenting style is commonly categorized into four different types (i.e., authoritative, authoritarian, neglectful, or indulgent/permissive) on the basis of two broad dimensions, parental control/demandingness and parental warmth/responsiveness (e.g., see Dallaire & Weinraub, 2005; Darling & Steinberg, 1993; Spera, 2005); see Figure 4.1. Parental control refers to the level of demands imposed on the child by the parents; high parental control is associated with high expectations of the child, high levels of monitoring, and with parents who provide discipline to their child when needed. Parental warmth refers to the level of responsiveness and supportiveness provided to the child by the parents; high parental warmth is associated with accepting, contingent, sensitive and involved parents and child-centered interactions. Previous research supports relative consistency in parenting style across time points (Dallaire & Weinraub, 2005).

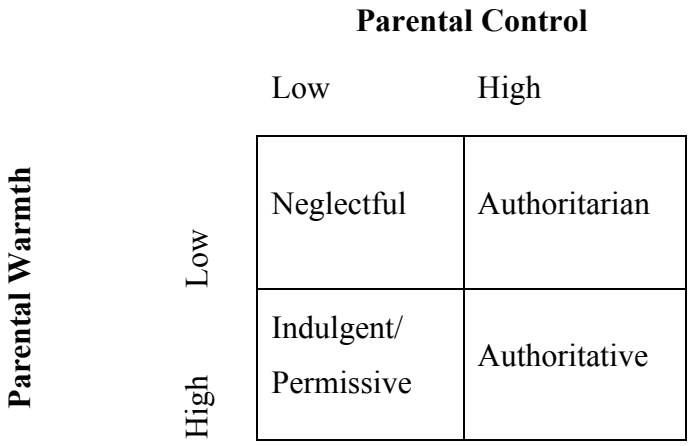


Figure 4.1. Parenting Style.

Parenting factors such as parental linguistic input (usually maternal; e.g., quantity, quality) and responsivity to the child are often studied within the field of CSD (Rogers et al., 2015); however, such studies are often limited by underspecified causal mechanisms. Parental typologies (i.e., authoritative, authoritarian, neglectful, or indulgent/permissive), although not as commonly studied within CSD, have been widely studied in related fields such as child development and psychology. In regards to parenting style dimensions, high parental warmth has been positively associated with communication development (e.g., Tamis-Lemonda, Bornstein, & Baumwell, 2001) and with school readiness more broadly (e.g., Landry, Smith, Swank, Assel, & Vellet, 2001) (which measures a child's cognitive, socioemotional, and attentional skills; Duncan et al., 2007),

at least in regard to certain cultural-linguistic groups. In contrast, high parental negative affect, negative control, and coercive discipline strategies (e.g., hitting, yelling) have been associated with socioemotional difficulties and lower academic achievement, particularly for European American children (e.g., Amato & Fowler, 2002; Dotterer, Iruka, & Pungello, 2012). Given language skills are closely associated with other cognitive skills such as executive function, attention, and memory, the association between parenting style dimensions and school readiness/academic achievement is also relevant to the development of language skills.

4.1.3.1.2 Potential mechanisms. Parenting style may influence language development via two main mechanisms, which may not be mutually exclusive. Parenting can influence the child's physiological response to stress which can alter brain development (Ben-Dat Fisher et al., 2007; De Bellis, 2001; Granger et al., 1998). A positive parent-child interaction, consisting of warm and nurturing parenting, might mitigate the impact of stress on a child's mental, cognitive and physical wellbeing, thereby increasing the child's receptivity and ability to learn. For example, Landry and colleagues (2001) reported that preschool age children, particularly preterm children, with consistently responsive mothers made faster cognitive gains than their peers with less responsive mothers. It is also possible for children to influence parental responsiveness, and children's level of receptivity may also differ. For example, children who have difficulty thriving may potentially increase parental stress and as a result decrease parental responsiveness. On the other hand, a negative parent-child interaction may negatively influence learning and memory. Chronic childhood stress, leading to high levels of cortisol, has been associated with various types of physical and mental illnesses and childhood neglect is known to increase adult risk for morbidity and mortality (e.g., Carroll et al., 2013).

Another mechanism is the potential for parenting style to make the child more or less receptive to socialization. Parent-child interaction plays a key role in language learning as parents are highly influential in the nature of children's learning environments and often in a position to scaffold the child's language learning, consistent with a Vygotskian view of human development. Accordingly, disruptive or negative patterns of parent-child interaction may lead to fewer scaffolded interactions and contribute to more pervasive negative socioemotional outcomes across contexts.

4.1.3.1.3 Empirical research. Although critical cross-cultural differences have been noted in the associations between parenting styles and child outcomes, particularly in relation to race/ethnicity, culture, and socioeconomic status (Baldwin, Baldwin, & Cole, 1990; Dotterer et al.,

2012; Garcia Coll, 1990; Lamborn, Dornbusch, & Steinberg, 1996; Pungello, Iruka, Dotterer, Mills-Koonce, & Reznick, 2009; Spera, 2005), other studies indicate cross-cultural validity for a core set of parenting styles associated with positive child outcomes, including language. Specifically, warm, structured, responsive parent-child interactions, and avoidance of coercive discipline strategies have been found to be associated with positive outcomes across diverse types of families (Amato & Fowler, 2002; Bradley & Corwyn, 2000; Connell & Prinz, 2002).

Research examining the association between parent-child interactions and children's later language development is consistent with findings linked to a common core of parenting styles leading to positive outcomes. High parental warmth (e.g., emotional responsiveness, sensitivity) and avoidance of coercive discipline strategies are associated with better early language knowledge and literacy development (Connell & Prinz, 2002; Dodici, Draper, & Peterson, 2003; Mistry, Biesanz, Taylor, Burchinal, & Cox, 2004; Pianta, 1997; Pianta, Nimetz, & Bennett, 1997; Tamis-Lemonda et al., 2001). In a longitudinal study with 40 mother-child European-American dyads (Mean age of children = 9.5 ms and 13.7 ms), increased maternal responsiveness, measured by coding of videotaped interactions, was positively associated with expressive language milestones, based on maternal report (e.g., Early Language Inventory, MacArthur-Bates Communicative Development Inventory) (Tamis-Lemonda et al., 2001). Similarly, in a study with 47 African American kindergarteners (Mean age = 5.4 years), structured and emotionally responsive interactions, measured by coding of parent-child videotaped interactions, were associated with higher scores of overall communication skills and receptive communication skills as measured by the Battelle developmental inventory, with negative control associated with decreased receptive communication skills (Connell & Prinz, 2002). The effect of specific parenting style dimensions on school-age language development is less clear.

4.1.3.2 Reading exposure.

4.1.3.2.1 Background. The association between reading and overall cognitive development has been widely studied within the fields of psychology, education, and communication sciences and disorders. Factors such as early home literacy environment (e.g., home reading exposure[23]), early language skills, early reading skills, reading interest, and accessibility to literacy resources have

[23] Home literacy environment includes reading exposure as well as exposure to videotapes, and engaging in sociodramatic play, singing, drawing and writing.

been shown to have positive associations with language and literacy development as well as later academic success (Agostin & Bain, 1997; Burgess, Hecht, & Lonigan, 2002; Guthrie & Wigfield, 2000; McCardle et al., 2001; Petrill, Deater-Deckard, Schatschneider, & Davis, 2005; Scarborough & Dobrich, 1994; van Steensel, 2006). Home reading exposure, as construed here, refers to solo- or joint-book reading or listening to digital or print texts which may include print books, e-books, and picture-books in the home environment. Longitudinal studies demonstrate positive correlations between early home literacy practices, including reading exposure, and breadth of early childhood vocabularies (Hart et al., 2009; van Steensel, 2006). In addition to positive correlational data, intervention studies indicate joint book-reading can facilitate the development of language skills in the preschool years, although evidence for long-term benefits is limited (see Scarborough & Dobrich, 1994 for a review).

4.1.3.2.2 Potential mechanisms. Reading exposure may influence language development via two main mechanisms. One potential mechanism, embedded within a Vygotskian framework of development, is that engaging in reading activities, with social guidance from a more experienced partner, can facilitate language development. Joint reading with a caregiver provides direct opportunities to scaffold language-learning with a more knowledgeable person who is able to provide linguistic input that is appropriate for the child's linguistic level and has the potential of providing indirect and direct language instruction. As such, joint reading provides children with the opportunity to witness language-rich models, and to practice language use. Research supports the hypothesis that reading exposure may promote language development (Scarborough & Dobrich, 1994; Sénéchal, LeFevere, Thomas, & Daley, 1998). Reading exposure may also act as a general indicator of parental investment in learning, at least for some cultural-linguistic groups (Hart et al., 2009).

A second and related potential mechanism linking reading exposure to language development is the direct opportunity it provides to develop linguistic resources. Reading is a language-based skill, with partially overlapping brain networks (e.g., perisylvian region) (Schlaggar & McCandliss, 2007), specifically as it relates to networks supporting phonological, syntactic, and semantic skills. The overlapping skills between reading and language are supported by common shared environmental and genetic influences (Harlaar, Hayiou-Thomas, Dale, & Plomin, 2008). Whereas literacy acquisition (e.g., decoding) is thought to draw primarily on children's pre-existing phonological (e.g. phoneme segmentation, phonological decoding), and syntactic skills

(i.e., knowledge of the structure of language), later stages of reading development are viewed as a direct support for language learning, with syntactic and semantic (e.g., vocabulary) skills playing a role in reading comprehension (Cromley & Azevedo, 2007; Harlaar et al., 2008; Scarborough & Dobrich, 1994). The association between reading and language is supported by research indicating children's early language skills have been found to be a good predictor of reading outcomes later in development (Catts, Fey, Zhang, & Tomblin, 1999; Elbro & Scarborough, 2004; Muter, Hulme, Snowling, & Stevenson, 2004; Scarborough, 2005). Moreover, school-age readers have been reported to use reading as a source for language learning (Cunningham, 2005). In "The stages of reading development", Chall (1983) describes this stage, between the ages of 9 – 13 as the stage in which children read to learn new knowledge.

4.1.3.2.3 Empirical research. The present section will highlight a correlational study and an intervention study examining the role of reading in young children, followed by a behavioral genetic study in school-age children—all of which emphasize the role of reading throughout language development. Correlational studies have provided support for the positive association between reading and language. In young children ages 8-24 months, reading to them at least once a day, instead of less frequently, was significantly associated with larger vocabulary, based on regression analysis, as measured by the short-form of the MacArthur-Bates Communicative Development Inventory (CDI), a well-validated standard measure of child language development (Fenson et al., 2000). Additionally, parents telling stories to their children at least once a day, another type of literacy activity, was associated with higher CDI scores in both infants and toddlers, albeit only the toddler group (17-24 ms) reached statistical significance (Zimmerman, Christakis, & Meltzoff, 2007). Reading exposure has been shown to account for variance in children's receptive and expressive vocabulary even after controlling for parental education, parental literacy skills, and children's analytic intelligence (Sénéchal, LeFevre, Hudson, & Lawson, 1996).

Intervention studies that use reading exposure to facilitate language development further support positive outcomes in receptive and expressive language (e.g., Hargrave & Sénéchal, 2000; Isbell, Sobol, Lindauer, & Lowrance, 2004; Newman, 1996). Hargrave and Sénéchal (2000) conducted a storybook reading intervention in 36 preschoolers with poor expressive vocabulary skills in which they manipulated the frequency and nature of book-reading. The experimental group participated in active story book-reading in which children were encouraged to actively

participate, and teachers/parents were encouraged to use questioning techniques, provide feedback to the children, adapt their reading style to the child's linguistic level, and have fun during the reading activity. The active control group participated in 'regular' (passive) book reading in which teachers were encouraged to read to the children as they usually did. For children in both groups, teachers were given 10 books to read, and each book was read twice for the children to benefit from repeated exposure to the books; parents were given 4 books to read at home with their child throughout the intervention and were instructed to read each book five times. At the end of the four-week study, children in both groups showed improvements in expressive vocabulary introduced in the books, with children in the active reading group showing significantly greater improvements in vocabulary introduced in the books and the Expressive One Word Picture Vocabulary Test – Revised compared with the children in the regular reading group.

Finally, findings from the WRRMP have supported the influence of shared home literacy environment on early expressive vocabulary in school-age children using twin methodology. Specifically, Hart and colleagues (2009) utilized measures of home literacy environment, assessed via a parental questionnaire of reading behaviors in the home, and expressive vocabulary, assessed via the Boston Naming Test, from 314 twin-pairs at three different time points (mean ages = 6, 7, 8). Using structural equation modeling, results suggested that the home literacy environment accounted for 6-10% of the shared environmental variance in children's expressive vocabulary at school-age after controlling for genetic influences. In contrast to the present study, Hart et al., (2009) relied on an untraditional measure of child vocabulary; the Boston Naming Test is a measure of word-retrieval normed on aphasic and non-aphasic adults. Additionally, the home literacy environment was assessed as a shared environmental variable and did not examine twin-specific differences in reading exposure within the home (i.e., nonshared environment).

***

In sum, the proposed study will examine the potential influence of a specified set of environmental factors (i.e., birthweight, breastfeeding, home reading exposure, and parenting) on children's language outcomes at school-age using a genetically-sensitive design. This study is an explicit extension of prior work which has highlighted the need to study a broader range of environmental influences on child language development and to address the interplay between genetic and environmental influences (Rogers et al., 2015), particularly as related to nonshared environmental effects (i.e., DeThorne et al., 2012; Harlaar et al., 2016). Although quasi

experimental, this study provides specific advantages including the ability to a) control for genetics and shared environmental effects, b) examine associations with language across two specific time points, and c) incorporate both standardized tests and narrative language measures. To the best of my knowledge, the present study offers the first application of the MZ twin differences method to understanding specific NSE influences on children's language development. The specific research questions are as follows:

1. To what extent are MZ differences in birthweight, breastfeeding, home reading exposure, and parenting associated with MZ differences in Formal and Productive language at mean age 10 (i.e., HV5)?
2. To what extent do associations between discordance in candidate environmental measures and language outcomes replicate across time points (i.e. HV5 and HV7)?
3. To what extent does the association of discordance in candidate environmental measures and language outcome measures differ based on the extent of language discordance?

4.2 Method

4.2.1 Participants. Participants included monozygotic same-sex/gender[24] twins (56-58% girls) drawn from the Western Reserve Reading and Math Project (WRRMP; Petrill, Deater-Deckard, Thompson, DeThorne, & Schatschneider, 2006), a longitudinal twin study of reading development and related cognitive skills. All participants were originally recruited to the larger project from throughout Ohio during kindergarten/first grade via media advertisements, school nominations, Ohio birth records, and mothers-of-twins clubs, and followed through annual home visits until at least 5th grade. Annual home visits (HV) lasted approximately 2.5 hours and included standardized measures, narrative language samples, video-taped parent-child interactions, and parental questionnaires. Each twin within a pair was simultaneously assessed by different examiners to avoid inflation of twin similarity. DNA testing from buccal swabs, or a parent questionnaire of twin similarity reported to be 95% accurate (Goldsmith, 1991), was used to assign zygosity.

The present study focused on parenting and reading data from HV2, and language data from HV5 and HV7, which corresponds with 1st grade, 3rd grade, and 5th grade respectively. Data on

[24] Boy/girl distinction is based on binary parent identification of "sex" with options "girl" or "boy", thereby conflating sex and gender.

birthweight and breastfeeding were gathered at the point of initial enrollment in the study. To be included in the present analyses, MZ twin pairs needed to have complete data in terms of zygosity, age, and sex. In addition, both twins within a pair needed to have data for at least one language measure and one environmental variable. These selection criteria led to 230 twins who had language data at HV5 (*M* age = 9.83, *SD* = .95) as well as birthweight and/or breastfeeding data, and 162 twins who had language data at HV5 as well as parenting and/or reading data that was derived from HV2 (*M* age = 7.11, *SD* = .65). When we gathered language data at HV7, there were 216 twins who had language data (*M* age = 12.22, *SD* = 1.24) as well as birthweight and/or breastfeeding data, and 154 twins who had language data as well parenting and/or reading data at HV2 (*M* age = 7.13, *SD* = .70). When considering missing data for any individual variable, analyses included anywhere between 114 to 230 twins depending on the home visit and variables under consideration.

Based on results from the Speech-Language Survey (DeThorne et al., 2006), between 8-11% of the MZ sample was receiving speech-language pathology services when they entered the study. Consistent with the larger WRRMP database, the selected MZ sample primarily consisted of children of White/European Americans (93-95% compared to 83% in Ohio census data, 2010) with a high school education or higher (99-100% compared to 89% in Ohio census data, 2010). Parental education in the present sample was similar for mothers and fathers with 83-86% pursuing post high-school education: 25-26% some college/career school or 2-year degree, 30-35% 4-year degree, 3-5% some post-graduate education, and 22-25% completed post-graduate/professional school[25]. The majority of households in the MZ sample were opposite-sex[26] (100% of those who reported having a partner, with 1-3% not reported), two-parent households (95-97%, with 0-1% not reported).

4.2.2 Language outcome measures. The large variability in etiological influences on language development has been attributed in part to different forms of measurement (e.g., DeThorne et al., 2008; Harlaar et al., 2016). To account for this, language outcome measures were examined from both standardized language measures and narrative language measures at HV5 and HV7.

[25] Parental race and education percentages are based on data collected when the families first entered the WRRMP study, with 1-3% not reported.
[26] Parent sex is based on binary identification of "sex" with options "female" or "male".

4.2.2.1 Standardized language measures. Consistent with prior work (DeThorne et al. 2012; Harlaar et al., 2016), this study included the Oral Narration score from the Test of Narrative Language (TNL; Gillam & Pearson, 2004), and three subtests from the Clinical Evaluation of Language Fundamentals-Fourth Edition (CELF-4) (Semel, Wiig, & Secord, 2003) (recalling sentences, understanding spoken paragraphs, and word classes receptive and expressive), each detailed below.

4.2.2.1.1 TNL. This study used the Oral Narration score from the TNL (raw score: 0-90, the higher the score the better), which is based on 3 expressive tasks: 1) listening to and retelling a story about a trip to a fast-food restaurant without visual support, 2) telling a story related to a sequence of five drawings of a boy late for school, and 3) a picture-elicitation task in which participants needed to tell a story based on a picture of two children who witness a family of aliens landing at a park. The internal consistency (coefficient alpha) for the Oral Narration score for children ages 6-11 ranged from .88 to .87. Two-week test-retest reliability for a sample of 27 children ages 5-10 years was .80 (uncorrected).

4.2.2.1.2 CELF-4 subtests.

Recalling sentences. This subtest focuses on assessment of language memory and expression by evaluating the child's ability to a) listen to and attend to spoken sentences of increasing length and complexity (e.g. *Does Mr. Lopez teach reading?*), b) maintain this information in working memory, and c) repeat the sentences verbatim. Errors include omission, repetition, addition, transposition, and substitution of words or parts of words. Raw scores range from 0-96; the higher the score the better. Subtest internal consistency (coefficient alpha) for children ages 6-13 ranged from .86 to .91. Test-retest reliability (Range = 7-35 days, M = 16 days, n = 320) for a sample of 252 children ages 6-13 years ranged from .87 to .92 (uncorrected).

Understanding spoken paragraphs. This subtest is intended to assess receptive language by evaluating the child's ability to a) listen and attend to three spoken paragraphs of increasing length and complexity and b) answer 15 questions about the content regarding main ideas, details, sequence of events, inferences, and predictions. Raw scores range from 0-15; the higher the score the better. Subtest internal consistency (coefficient alpha) for children ages 6-13 ranged from .62 to .74. Test-retest reliability (Range = 7-35 days, M = 16 days, n = 320) for a sample of 252 children ages 6-13 years ranged from .51 to .87 (uncorrected).

Word classes receptive and expressive. This subtest was designed to assess receptive and expressive language by evaluating the child's ability to understand and explain relationships in the meanings of associated words. The child listens to a set of four words, selects two words that are related, and explains the association between the words (e.g., Examiner*: tell me the two words that go together: fish, milk, fin, spider.* Child*: fish & fin.* Examiner*: How are the words fish and fin related?*). The child is given credit for the expressive component if the receptive component is correct. Raw scores range from 0-42 for children ages 5-8 and 0-48 for children ages 9-21; the higher the score the better. Subtest internal consistency (coefficient alpha) for children ages 6-13 ranged from .85 to .91. Test-retest reliability (Range = 7-35 days, *M* = 16 days, *n* = 320) for a sample of 252 children ages 6-13 years ranged from .68 to .90 (uncorrected).

4.2.2.2 Language sample measures. Narrative language samples were collected via the three expressive tasks of the Test of Narrative Language and an additional picture-elicitation task that included telling a story based on a picture taken from the Test of Language Competence-Expanded Edition (TLC-E; Wiig & Secord, 1989). The language sample was audio-recorded on a compact flash card using a Marantz recorder and transcribed by research assistants in the Child Language and Literacy Laboratory at the University of Illinois, based on the Systematic Analysis of Language Transcripts (SALT) conventions (Miller, Iglesias, & Nockerts, 2004). Each twin in a pair was transcribed by a different research assistant, and transcribers were naive to zygosity, to avoid potential inflation of MZ twin similarity. Research assistants achieved 85% point-by-point agreement with an experienced transcriber during training on utterance boundaries and individual morphemes. Transcription reliability checks confirmed agreement remained between .67 – 1.00 for HV5 and .84 – 1.00 for HV7 (see Harlaar et al., 2016 for details). As is recommended for school-age children, the sample was segmented into communication units (C-units), separating independent clauses joined by conjoining conjunctions (i.e., and, but, or) into distinct utterances to avoid inflating the length of utterances (Loban, 1976; Nippold, 1998). Repeated or reformulated segments were not included in the linguistic analysis. The MZ language samples from HV5 and HV7 averaged 62 (*SD* = 37, Range = 20 – 297) and 60 (*SD* = 24, Range = 27 – 236) C-units respectively. There is support for the validity and reliability of language sample measures up to age 13 (Heilmann, Nockerts, & Miller, 2010; Miller, Freiberg, Holland, & Reeves, 1992; Miller et al., 2005; Rice, Redmond, & Hoffman, 2006). Additionally, Harlaar and colleagues (2016) provided support of social validity and convergent validity of language sample measures for a

larger WRRMP sample of children at HV5 and HV7. Consistent with Harlaar and colleagues (2016), specific language sample measures for the present analyses included mean length of utterance (MLU), number of total words (NTW) and number of different words (NDW), each delineated as follows.

4.2.2.2.1 MLU. Mean length of utterance is a measure of a child's productive abilities. It was calculated by measuring the average length of all complete and intelligible C-units in morphemes. Consistent with established conventions and prior procedures (DeThorne et al., 2008, 2012; Harlaar et al., 2016; Miller et al., 2005), bound inflectional morphemes (e.g. past tense -ed, plural s, present progressive -ing), but not derivational morphemes (e.g. dis- in dishonest, -be in befriend), were counted as separate morphemes.

4.2.2.2.2 NTW. Number of total words is a measure of a child's semantic productivity. To decrease the influence of volubility and maximize the available data, a frequency count of all root word tokens within the first 30 complete and intelligible C-units was calculated (Harlaar et al., 2016).

4.2.2.2.3 NDW. Number of different words is a measure of a child's vocabulary diversity. It was derived as a frequency count of different word tokens within the first 30 complete and intelligible C-units (cf. DeThorne, Deater-Deckard, Mahurin-Smith, Coletto, & Petrill, 2011; Harlaar et al., 2016; Hutchins, Brannick, Bryant, & Silliman, 2005).

4.2.2.3 Language outcomes factor analysis. Latent factor analysis explores underlying hidden or unobservable factors in a set of measured variables by reducing the measured variables to fewer latent factors that share a common variance and are free of measure-specific variance and error. Based on prior work from WRRMP (e.g., DeThorne et al., 2008; Harlaar et al., 2016), language measures were expected to load on two distinct latent factors based on form of measurement: standardized language measures (i.e., Formal language) versus language sample measures (i.e., Productive language). In the present analyses, Formal and Productive language composites were computed to maximize available data by standardizing ($(X - \mu) / \sigma$), averaging, and re-standardizing accordingly ($M = 0$, $\sigma = 1$). Language composites were scaled so that a higher score indicates more advanced Formal or Productive language skills.

4.2.3 Environmental measures. Environmental influences in the present study are defined broadly as non-genetic influences (e.g., extent of breastfeeding) with a focus on child-specific

variables that can have nonshared effects. NSE influences will be derived via difference scores in twin-specific measures, thereby representing differential experiences in identical twins reared within a shared environment. Consistent with prior literature, environmental measures are delineated as a) perinatal and early postnatal measures, and b) social and linguistic measures (see Table 4.1 for a summary of the measures included).

4.2.3.1 Perinatal and early postnatal measures. Consistent with prior studies that assess birthweight and feeding practices based on parent recall (e.g., Mahurin-Smith & Ambrose, 2013; Mahurin-Smith et al., 2014) these variables were assessed via an intake questionnaire completed by the parents (usually mothers) upon enrollment to the WRRMP. Specifically, birthweight was based on a single question, and questions regarding feeding practices focused on initiation and end date for breastfeeding[27] and formula feeding (specific questions used for analyses are included in Table 4.1). Parent-reported birthweight has been found to have evidence of high validity (r = .97; ICC = .94) and reliability (r = 97; ICC = .93) when compared to hospital-recorded birthweight (Adegboye & Heitmann, 2008; see also Pyles, Stolz, & Macfarlane, 1935). Maternal recall of breastfeeding onset and duration have also been found to have evidence of high validity and reliability compared to medical records or prospectively collected data, with validity and reliability for recall decreasing with time (see Li, Scanlon, & Serdula, 2005 for a review). For example, when compared to medical records 79% of mothers (n = 64) recalled the breastfeeding duration of children 1-10 yrs within one month, and 95% within 2 months (Eaton-Evans & Dugdale, 1986). Similarly, when mothers (n = 146) were asked to recall breastfeeding duration at two different time points, two years apart (M = 51 and 53 years post- birth), the Spearman rank correlation coefficient was r = .86 (Tomeo et al., 1999). In the present analyses, birthweight and breastfeeding duration were scaled so that a higher score indicates increased birthweight or breastfeeding duration accordingly.

4.2.3.2 Social and linguistic measures.

4.2.3.2.1 Parenting. Assessment of parenting comes from two main sources: a) parent's self-reported measures of feelings toward each twin via the Parent Feelings Questionnaire (PFQ), and b) observer ratings of in-home videotaped dyadic interactions using the Parent-Child Interaction

[27] Mothers reported breastfeeding directly from the breast or tube/bottle-feeding expressed milk.

System (PARCHISY). Both forms of assessment provide an overall positivity and negativity index, yielding four total measures of parenting, each scaled so that a higher score indicates more parental positivity or negativity.

Table 4.1. *Environmental Measures.*

Perinatal & Early Postnatal Measures	Question(s)/Method	Scale/Unit
Birthweight	1. What were the weights of the twins at birth?	Kilogram
Extent of breastfeeding	1. Did you (twins' mother) breastfeed the twins? 2. If yes, when did you begin? 3. And when did you stop?	Months (Bf duration)
Social & Linguistic Measures	**Question(s)/Method**	**Scale/Unit**
Parenting	1. Parent Feelings Questionnaire: a. Scale A: consisted of statements such as "I enjoy hugging and cuddling with [this child]" and "Sometimes I find it difficult to be around [this child]". b. Scale B: consisted of 10 emotions toward each child (e.g., happy, sad, angry, excited, hostile). 2. Parent-Child Interaction System: video-taped sessions of parent-child interactions were used to rate parent positive/ negative affect, positive/ negative content/ control and responsiveness.	1. Parent Feelings Questionnaire: each scale had a positivity and negativity subscale. a. Scale A: used two 5-point Likert-type scales (1= definitely untrue for me, 5= definitely true for me). b. Scale B: used two 10-point frequency scales (1= never, 10= always). 2. Parent-Child Interaction System: five 7-point Likert-type scales
Home Reading Exposure	1. A) Some parents have the opportunity to read with their children. Who reads to [this child]? B) On average, how often do these people read to [this child]? 2. In a typical week, how many times do you and [this child] read books together? 3. How often does [this child] ask you to read books to him/her? 4. Describe how much [this child] enjoys being read to?	1. A) 0-1: no one reads to [this child], someone reads to [this child] B) 1-3 for three possible caregivers: daily, several times a week, weekly or less 2. 1-5: more than 3 times per day, once a day, once a week, once a month, almost never 3. 1-5: more than 3 times per day, once a day, once a week, once a month, almost never 4. 1-4: very much, somewhat, a little, not at all

PFQ. Parents (usually mothers) within WRRMP completed the Parent Feelings Questionnaire (Deater-Deckard, 1996, 2000), thereby providing parental positive and negative feelings towards each child via two scales, each with a positivity and a negativity subscale. Although the PFQ focuses on parental feelings, it also encompasses statements targeting parental behaviors and communication style. Scale A includes 24 statements such as "I usually make an effort to praise [this child] for good behavior" and "Sometimes I am not happy about my relationship with [this child]", and uses a 5-point Likert-type scale (1 = definitely untrue for me, 5 = definitely true for me). Scale B consists of 10 positive (i.e., amused, excited, happy, joyful, and proud) and negative (i.e., sad, angry, hostile, frustrated, furious) emotions towards each child, and uses a 10-point frequency scale (1= never, 10= always). When possible, data were averaged across mothers and fathers to yield more reliable composite scores and for sex/gender inclusion. Consistent with prior studies (Deater-Deckard, 2000; Wang, 2013), the positivity subscales (16 total items) were used to calculate an overall score of parent self-reported positivity; the negativity subscales (18 total items) were used to calculate an overall score of parent self-reported negativity. An overall positivity and negativity z-score was computed by standardizing and then averaging the scores from the two positivity subscales, and doing the same for the two negativity subscales, and standardizing the averaged scores again (M = 0, σ = 1). Previous studies using the PFQ have reported Cronbach's alpha coefficients from .67 - .87 for the positivity scales and .80 - .93 for the negativity scales (Deater-Deckard, 2000; Wang, 2013). The PFQ positivity and negativity scales have shown to be moderately to highly correlated (r = .40 - .70) with other parental affect measures (Deater-Deckard, 2000), including the warmth and negativity scales of the Parent Report (Dibble & Cohen, 1974).

PARCHISY. The observer ratings of parent interactions focused on two videotaped 10-min cooperative-play tasks between the primary parental caregiver (usually mothers) and each twin separately. Dyads were observed using an Etch-A-Sketch drawing toy (adapted from Stevenson-Hinde & Shouldice, 1995) with two knobs: one knob for drawing vertically, and another knob for drawing horizontally (see Appendix B). The dyads were asked to work together to cooperatively draw a house (see Appendix B). In the second structured interaction, dyads were observed using a Marble Maze game with two knobs: one knob for tilting, and one knob for rolling (see Appendix B). The dyads had to work together to navigate a marble by cooperatively tilting and rolling the

wooden maze. To encourage cooperation, the parent and the child were each assigned a knob, and were told not to use each other's knobs.

Trained research assistants in the Gene-Environment Processes laboratory at the University of Oregon completed ratings of the two videotaped tasks for each parent-child dyad using the Parent-Child Interaction System, PARCHISY (Deater-Deckard, 2000; Deater-Deckard, Pylas, & Petrill, 1997). PARCHISY is a global behavioral coding scheme designed for use with children ages 3-12 that has been widely used with various populations to examine naturalistic interactions within research settings. To avoid potential rater bias, different observers rated interactions between each twin within a pair and their mother, and observers were blind to zygosity and to the interactions between the parent and the other twin. All observers reached Cronbach's alpha > .75 during training and maintained this level of agreement. Observers rated child-specific parental affect (positive/negative), content/control (positive/negative), and responsiveness by completing five 7-point Likert-type scales (from low to high frequency). Positive affect was defined as expressing emotions with positive valence such as smiling and laughing. Negative affect was defined as expressing emotions with negative valence such as frowning and a cold/harsh voice. Positive content/control was defined as the use of praise, explanation, and open-ended questions. Negative content/control was defined as the use of criticism or intentional physical control of the child's hand/arm/body or the child's knob. Parental responsiveness was defined as the degree and immediacy with which the parent responded to the child's questions, comments, and behaviors. Predictive validity of the measure is supported by associations with various child outcomes, including socio-emotional adjustment based on parent's reports (e.g., see Deater-Deckard et al., 2001). Deater-Deckard (2000) reports Cronbach's alpha coefficients between .86 (positive content/control) to perfect reliability with a one point allowance (negative affect, negative content/control). The Etch-a-Sketch and Marble Maze interactions were rated separately and the scores were averaged across tasks. Consistent with Mullineaux and colleagues (2009), a WRRMP longitudinal study, a positivity (i.e., positive affect, content/control, and responsiveness) and negativity (i.e., negative affect and content/control) composite were computed. To assess the internal consistency and dimensionality of the scales, Cronbach's alpha and principal component analysis were used.

4.2.3.2.2 Home reading exposure. Home reading exposure (HRE) was examined via four specific items addressing exposure and interest in book reading taken from the WRRMP parental

questionnaire 2 at HV2 (see Table 4.1): Item 1 was originally derived from Griffin and Morrison's (1997) Home Literacy Environment Scale; items 2-4 were originally derived from Petrill and colleagues (2005). In regard to scale unidimensionality and reliability, Petrill and colleagues (2005) reported that the Educational Progress Survey, consisting of a largely overlapping set of questions, loaded on a single factor and demonstrated a Cronbach's alpha of .90. Specific to the Home Reading Exposure questionnaire, Item 1.a and 1.b were merged into a 1-10 scale (1 = no one reads to this child, 10 = three different caregivers read to this child daily). When possible, for the present analyses, data were averaged across mothers and fathers to yield more reliable composite scores and for sex/gender inclusion. Similar to Petrill and colleagues (2005), the HRE scale was evaluated for unidimensionality and internal consistency. A composite score was computed by standardizing each item, averaging the scores across the items for each child, and standardizing the averaged scores again ($M = 0$, $\sigma = 1$). All Home Reading Exposure items were scaled so that a higher score indicates increased reading exposure.

4.2.4 Analyses. To examine nonshared environmental influences on language development, an MZ-difference method was used (Mullineaux et al., 2009; Pike et al., 1996). The MZ-difference method is a simple, yet sensitive method that posits differences between MZ twins can be used as an index of nonshared environmental effects. Since MZ twins share approximately all of their genes, NSE effects can be assessed independent of genetic and shared environmental confounds. Previous research has suggested the MZ-difference approach is more sensitive to NSE influence than the multivariate genetic approach (Pike et al., 1996). Although a sensitive measure, the MZ-difference method does not assess causality or directionality of the relationship between measured environments and measured language outcomes.

As is customary within the MZ-difference methodology, each twin pair was randomly assigned as Twin 1 or Twin 2 to control for confounding variance. Preliminary analyses included descriptive statistics, and monozygotic intrapair correlations for each language composite and environmental measure to assess the adequacy of the measure for use in NSE analyses (i.e., higher MZ intrapair correlations would suggest lower overall NSEs). To address the first question, examining the degree to which differences in environmental factors relate to differences in Formal and Productive language, relative differences for each environmental measure (e.g., birthweight) were calculated (Twin 2 environmental score – Twin 1 environmental score) and correlated with relative differences for each language measure (i.e., Twin 2 language score – Twin 1 language score) at

mean age 10 (HV5). A correlation of ±1.0 would indicate a perfect linear association between the discordance for the language outcome measure and the discordance for the environmental measure, whereas a correlation of 0.0 would indicate no linear relationship. Additionally, to address the second question, examining the extent to which findings replicate across time points, analyses were replicated in relation to language development at mean age 12 (HV7).

Finally, to address the third question, exploring NSE effects at the extremes of the language distributions, simple linear quantile regression conditional on language differences was used. Quantile regression, developed by Koenker and Bassett (1978), offers an alternative approach to conditional means regression (e.g., ordinary least squares), by examining the relationship between an independent and a dependent variable conditional on quantiles across the distribution of the dependent variable. In relation to the present study, quantile regression allows for the estimation of the relation between language difference scores and environmental difference scores at the extremes of the language distribution using the full data set. More specifically, quantile regression uses the full data set to estimate regression coefficients by using asymmetric weighting of positive and negative residuals as a function of the quantile of interest (for additional details on quantile regression see Cade & Noon, 2003; Koenker, 2005; Koenker & Hallock, 2001; Petscher & Logan, 2014). Difference scores were re-calculated for quantile regression so that the score of the twin with the lower language was subtracted from the score of the twin with the higher language. This was done to ensure lower quantiles represented the least degree of language difference between twin pairs and higher quantiles represented the largest degree of difference between twin pairs. The order used to calculate language difference scores was maintained when calculating differences in environments.

4.3. Results

4.3.1 Preliminary analyses. Descriptive statistics for each language measure at mean age 10 (HV5) and 12 (HV7), and each environmental variable, presented as raw scores, are summarized in Tables 4.2 and 4.3.

4.3.1.1 Language outcomes. In general, descriptive language outcomes were consistent with developmental expectations. Specifically, productive language scores increased slightly from HV5 to HV7 (Miller et al., 2005), and mean standardized scores from the CELF-4 subtests approximated normative means. In particular, standard scores from the sentence repetition task (i.e., Recalling

Sentences), a relatively sensitive marker of language impairment (e.g., Conti-Ramsden, Botting, & Faragher, 2001), indicated that the majority of the scores were within (71-72%) or above (22%) one standard deviation from the mean across time points. Only 6-7% of the scores were below one standard deviation from the mean, demonstrating the MZ sample primarily consisted of children with typical or above average language skills.

Consistent with prior work (e.g., DeThorne et al., 2008; Harlaar et al., 2016), the present data set revealed high correlations among productive language sample measures (.70 to .89 among MLU, NTW, and NDW), and moderate to high correlations among the standardized language measures (.31 to .64 among TNL, and the 3 CELF subtests) within both home visits (HV5 and HV7). In contrast, there were small to moderate correlations across language sample measures and the standardized language measures (.16 to .46). An oblique latent factor analysis confirmed language sample measures and standardized-test scores loaded onto two separate factors. Language sample measures loaded highly on the latent Productive language factor at HV5 and HV7 (>.80), whereas loadings of the standardized language measures on the Formal language factor were moderate to large (.48 - .98). Cronbach's alpha for the Productive language factor was .93 at both HV5 and HV7, indicating excellent internal consistency. Cronbach's alphas for the Formal language factor were .81 (HV5) and .75 (HV7), indicating good and acceptable internal consistency respectively. As expected, language composites were found to correlate substantially with extracted latent factor scores (Productive language: r = .97 to .99; Formal language: r = .91 to .99).

4.3.1.2 Environmental variables.

4.3.1.2.1 Birthweight. One pair of twins with parent-reported birthweight greater than three times the inter-quartile range was considered an extreme outlier and was removed from the sample. Consistent with twin pregnancies, on average twins weighed approximately 2.4 kg at birth. As expected, birthweight correlated highly with gestational age (HV5: r = 0.67, p = < .001; HV7: r = 0.70, p = < .001); gestational age did not differ within twin pairs.

4.3.1.2.2 Breastfeeding. Between 79% (HV5 sample) and 83% (HV7 sample) of mothers reported breastfeeding. Extent of breastfeeding was log-transformed (natural log) due to positive skew.

4.3.1.2.3 PFQ. Overall, parents reported experiencing higher positivity than negativity towards their child. On average, parents reported positive statements were definitely true for them, and rated the frequency with which they experienced positive emotions with a 9 out of 10 (1 = never, 10 = always). In contrast, on average parents reported negative statements were somewhat untrue for them, and rated the frequency with which they experienced negative emotions with a 2 out of 10. PFQ positivity scales A and B were reflected and log transformed (natural log) due to negative skew; variables were re-scaled so that the higher the score, the more parent self-reported positivity. In turn, PFQ negativity scales A and B were log transformed (natural log) due to positive skew. Cronbach's alpha coefficients ranged from .71 - .73 for the positivity scales and .78 - .80 for the negativity scales, indicating acceptable to good internal consistency.

4.3.1.2.4 PARCHISY. Similar to self-reported parenting, twins had higher positivity than negativity scores for parent-observed measures. For example, on parent-observed positivity measures parents scored on average a 3 in positive affect towards their child, a 4 in positive content/control, and a 6 in responsiveness (1-7 Likert-type scale), indicating a few/several instances of positive affect (i.e., smiling, laughing), moderate amounts of positive content/control (i.e., reliance on explicit directions with at least one instance of praise, explanation, or questioning), and substantial responsiveness in which the parent responded to most of the child's comments, questions, and behaviors, with no delay, with only 1 or 2 instances of non-responsiveness. In contrast, on average parents scored a 1 on parent-observed negative affect and negative content/control, indicating no negative affect and no negative content/control were displayed. The PARCHISY negativity composite was highly positively skewed, therefore an inverse (reciprocal) transformation was used. The composite was re-scaled so that the higher the score, the more parent-observed negativity. There were moderate to high correlations between parent-observed positivity measures (.48 to .53 among positive affect, positive content/control and responsiveness), and substantial correlations between parent-observed negativity measures (.65 among negative content/control and negative affect). In contrast, there were small to moderate correlations across positivity and negativity measures (- .18 to - .30). Principal Component Analysis confirmed the PARCHISY measures loaded highly onto two components accounting for 73 - 74% of the variance, a positivity component (eigenvalue = 2.44 - 2.45, percentage variance = 48.82 – 48.92%) and a negativity component (eigenvalue = 1.22 – 1.25, percentage variance = 24.48 – 24.97%). Cronbach's alpha for the positivity component ranged from .75 - .76; Cronbach's

Table 4.2. *Descriptive statistics for language measures at mean age 10 (HV5) and environmental variables.*

	N	Mean	SD	Min	Max
Language Outcomes at 12 (raw score)					
Standardized Language Measures					
CELF-4: Recalling sentences (0-96)	226	68.56	12.85	35.00	93.00
CELF-4: Understanding spoken paragraphs (0-15)	198	10.77	3.09	1.00	15.00
CELF-4: Word classes rec. and exp. (0-42/48)	226	19.23	6.26	1.00	35.00
TNL: Oral Narration Score (0-90)	209	52.08	11.23	17.00	78.00
Language-sample Measures					
Mean Length of C-Unit	172	9.28	1.35	5.79	14.26
Number of Total Words (30 C-units)	186	250.83	39.71	155.00	351.00
Number of Different Words (30 C-units)	186	121.78	15.43	79.00	173.00
Environmental Variables					
Birthweight (kilogram)	226	2.44	0.57	0.88	4.08
Breastfeeding (months)	216	4.28	6.19	0.00	36.00
Parenting at 7 (raw score)					
PFQ-Positivity					
A (5-point Likert-type scale: 11-55)	156	52.58	3.70	31.00	55.00
B (10-point frequency scale: 5-50)	156	43.76	4.40	25.00	50.00
PFQ-Negativity					
A (5-point Likert-type scale: 13-65)	156	25.70	9.03	13.00	50.00
B (10-point frequency scale: 5-50)	156	11.75	4.47	5.00	26.00
PARCHISY-Positivity (7-point Likert-type scales)					
Parent positive affect	140	3.08	1.11	1.00	7.00
Parent positive content/control	140	4.20	1.63	1.00	7.00
Parent responsiveness	140	6.05	0.93	3.00	7.00
PARCHISY-Negativity (7-point Likert-type scales)					
Parent negative affect	140	1.21	0.49	1.00	3.50
Parent negative content/control	140	1.25	0.60	1.00	5.50
Home Reading Exposure at 7					
1. Freq. anyone reads to child (1-10 scale)	154	5.12	1.63	1.00	10.00
2. Freq. parent & child read together (5-point Likert-type scale)	154	3.42	0.72	1.00	5.00
3. Freq. child elicits being read to (5-point Likert-type scale)	154	2.97	0.94	1.00	5.00
4. Extent child enjoys being read to (4-point Likert-type scale)	153	3.72	0.57	1.50	4.00

Note. CELF-4 = Clinical Evaluation of Language Fundamentals-Fourth Edition; Rec. = Receptive; Exp. = Expressive; C-unit = Communication unit; TNL = Test of Narrative Language; PFQ = Parent Feelings Questionnaire; PARCHISY = Parent-Child Interaction System; Freq. = Frequency. All measures/variables are scaled such that a higher score indicates more/higher language/environmental variable.

Table 4.3. *Descriptive statistics for language measures at mean age 12 (HV7) and environmental variables.*

	N	Mean	SD	Min	Max
Language Outcomes at 12 (raw score)					
Standardized Language Measures					
CELF-4: Recalling sentences (0-96)	216	76.98	11.22	42.00	96.00
CELF-4: Understanding spoken paragraphs (0-15)	184	12.41	2.27	5.00	15.00
CELF-4: Word classes rec. and exp. (0-48)	214	26.87	6.43	11.00	42.00
TNL: Oral Narration Score (0-90)	183	58.29	8.21	22.00	76.00
Language-sample Measures					
Mean Length of C-Unit	175	9.92	1.37	6.56	14.85
Number of Total Words (30 C-units)	177	267.29	41.94	163.00	426.00
Number of Different Words (30 C-units)	177	130.41	16.19	94.00	196.00
Environmental Variables					
Birthweight (kilogram)	210	2.43	0.60	0.88	4.08
Breastfeeding (months)	202	5.06	7.13	0.00	37.00
Parenting at 7 (raw score)					
PFQ-Positivity					
A (5-point Likert-type scale: 11-55)	146	52.33	3.76	31.00	55.00
B (10-point frequency scale: 5-50)	146	43.60	4.45	25.00	50.00
PFQ-Negativity					
A (5-point Likert-type scale: 13-65)	146	26.63	9.37	13.00	50.00
B (10-point frequency scale: 5-50)	146	12.22	4.83	5.00	26.50
PARCHISY-Positivity (7-point Likert-type scales)					
Parent positive affect	136	3.09	1.08	1.00	7.00
Parent positive content/control	136	4.20	1.63	1.00	7.00
Parent responsiveness	136	6.04	0.92	3.00	7.00
PARCHISY-Negativity (7-point Likert-type scales)					
Parent negative affect	136	1.21	0.50	1.00	3.50
Parent negative content/control	136	1.26	0.60	1.00	5.50
Home Reading Exposure at 7					
1. Freq. anyone reads to child (1-10 scale)	146	5.12	1.61	1.00	10.00
2. Freq. parent & child read together (5-point Likert-type scale)	144	3.38	0.72	1.00	5.00
3. Freq. child elicits being read to (5-point Likert-type scale)	144	2.99	0.88	1.00	4.00
4. Extent child enjoys being read to (4-point Likert-type scale)	145	3.71	0.57	1.50	4.00

Note. CELF-4 = Clinical Evaluation of Language Fundamentals-Fourth Edition; Rec. = Receptive; Exp. = Expressive; C-unit = Communication unit; TNL = Test of Narrative Language; PFQ = Parent Feelings Questionnaire; PARCHISY = Parent-Child Interaction System; Freq. = Frequency. All measures/variables are scaled such that a higher score indicates more/higher language/environmental variable.

alpha for the negativity component was .79 for both the subsample of children from HV5 (who also had PARCHISY data at HV2), and the subsample of children from HV7 (who also had PARCHISY data at HV2).

4.3.1.2.5 HRE. Item 1 of the Home Reading Exposure scale indicated that on average two caregivers (e.g., parents, siblings, sitters, relatives), or one caregiver and the child themselves read to each child several times a week, with a mean score of 5 in a 1-10 scale (1 = no one reads to this child, 10 = three different caregivers read to this child daily). The second and third items indicated on average parent(s) and child read together once a week, and children asked their parent(s) to read to them once a week, with mean scores of 3 on 5-point Likert-type scales (1 = almost never, 5 = more than 3 times per day). Finally, the fourth item indicated on average children enjoyed being read to very much, with a mean of 4 on a 4-point Likert-type scale. Inter-item correlations for the HRE scale were moderate to large (r = .33 to .62). The four items were found to load on one component using principal component analysis (eigenvalue = 2.49 to 2.52, percentage variance = 62.31 to 63.02%), indicating the scale is unidimensional. Cronbach's alpha ranged from .79 to .80, indicating acceptable to good internal consistency. The HRE composite was reflected and log transformed (natural log) due to negative skew. The composite was standardized again, and re-scaled so that higher scores indicate increased home reading exposure.

4.3.1.3 MZ correlations based on individual twin scores. All measures, with the exception of birthweight, were standardized either in the process of computing the composite or to ease interpretation of means and variances. Tables 4.4 and 4.5 list means, standard deviations, and MZ twin correlations based on individual twin scores, for language composites and environmental measures. Even though child sex/gender was controlled for in the primary analyses through use of same-sex/gender twins, we assessed mean differences between girls and boys for all environmental measures and language outcome measures using independent samples t-tests. As expected based on prior literature (e.g., UNICEF & WHO, 2004), boys were significantly heavier than girls for both the HV5 sample ($t(214.95) = 3.32$, $p = .001$) and the HV7 sample ($t(208) = 3.38$, $p = .001$). Sex/gender differences were also found for the Formal and Productive language composites at mean age 12 (Table 4.5), with girls having more advanced language skills than boys for Formal language ($t(214) = -2.02$, $p = .045$), and Productive language ($t(159.37) = -2.12$, $p = 0.04$). However, these differences vanished for the subsample of children who had language data at mean

Table 4.4. *Means, standard deviations, and MZ correlations (rMZ) based on individual twin scores for language composites at mean age 10 (HV5) and environmental measures.*

	No. of pairs	Mean	SD	rMZ
Language Outcomes at 10 (z- score)				
Formal language All	115	0.00	1.00	0.82**
Boys	48	-0.13	1.00	0.76**
Girls	67	0.09	0.99	0.86**
Productive language All	97	0.00	1.00	0.18*
Boys	43	-0.11	1.09	0.25+
Girls	54	0.09	0.92	0.10
Environmental Measures				
Birthweight All (kilogram)	113	2.44	0.57	0.83**
Boys	46	2.58	0.50	0.78**
Girls	67	2.34	0.59	0.85**
Breastfeeding All (months)	108	4.28	6.19	1.00**
Boys	46	4.07	4.70	1.00**
Girls	62	4.44	7.11	1.00**
[28]Parenting at 7 (z-score)				
PFQ-Positivity All	78	0.00	1.00	0.77**
Boys	30	-0.09	1.12	0.83**
Girls	48	0.06	0.92	0.73**
PFQ-Negativity All	78	0.00	1.00	0.80**
Boys	30	0.01	1.14	0.80**
Girls	48	-0.01	0.91	0.82**
PARCHISY-Positivity All	70	0.00	1.00	0.49**
Boys	27	-0.11	1.01	0.50**
Girls	43	0.07	0.99	0.50**
PARCHISY-Negativity All	70	0.00	1.00	0.51**
Boys	27	0.11	1.14	0.68**
Girls	43	-0.07	0.90	0.34*
Home Reading Exposure at 7 All (z-score)	78	0.00	1.00	0.90**
Boys	30	-0.16	1.09	0.95**
Girls	48	0.10	0.93	0.85**

Note. One-tailed *p*-values: +p=.05; *p<.05; **p<.01. Boy/girl distinction is based on binary parent identification of "sex" with options "girl" or "boy", thereby conflating sex and gender. PFQ = Parent Feelings Questionnaire; PARCHISY = Parent-Child Interaction System.

[28] Correlations between parenting measures (i.e., PFQ All & PARCHISY All) indicated a significant association between PFQ-Negativity All and PARCHISY-Negativity All (r(132) = .16, p = .03, one-tailed).

Table 4.5. *Means, standard deviations, and MZ correlations (rMZ) based on individual twin scores for language composites at mean age 12 (HV7) and environmental measures.*

	No. of pairs	Mean	SD	rMZ
Language Outcomes at 12 (z- score)				
Formal language All	108	0.00	1.00	0.80**
Boys	42	-0.17	1.01	0.77**
Girls	66	0.11	0.98	0.82**
Productive language All	89	0.00	1.00	0.22*
Boys	33	-0.20	0.86	0.13
Girls	56	0.12	1.06	0.24*
Environmental Measures				
Birthweight All (kilogram)	105	2.43	0.60	0.84**
Boys	40	2.60	0.54	0.77**
Girls	65	2.32	0.61	0.86**
Breastfeeding All (months)	101	5.06	7.13	1.00**
Boys	40	4.76	4.96	1.00**
Girls	61	5.25	8.27	1.00**
[29]Parenting at 7 (z-score)				
PFQ-Positivity All	73	0.00	1.00	0.76**
Boys	27	-0.07	1.15	0.82**
Girls	46	0.04	0.90	0.72**
PFQ-Negativity All	73	0.00	1.00	0.83**
Boys	27	-0.01	1.12	0.78**
Girls	46	0.01	0.93	0.88**
PARCHISY-Positivity All	68	0.00	1.00	0.52**
Boys	26	-0.11	0.97	0.58**
Girls	42	0.07	1.02	0.51**
PARCHISY-Negativity All	68	0.00	1.00	0.51**
Boys	26	0.11	1.14	0.70**
Girls	42	-0.07	0.90	0.33*
Home Reading Exposure at 7 All (z-score)	73	0.00	1.00	0.88**
Boys	27	-0.19	1.12	0.94**
Girls	46	0.11	0.91	0.82**

Note. One-tailed *p*-values: *$p<.05$; **$p<.01$. Boy/girl distinction is based on binary parent identification of "sex" with options "girl" or "boy", thereby conflating sex and gender. PFQ = Parent Feelings Questionnaire; PARCHISY = Parent-Child Interaction System.

[29] Correlations between parenting measures (i.e., PFQ All & PARCHISY All) indicated a significant association between PFQ- Positivity All and PARCHISY- Positivity All ($r(126) = .16$, $p = .04$, one-tailed).

age 12 (HV7) as well as parenting and/or reading data at mean age 7 (HV2). No other sex/gender differences were found [30].

The greater the MZ correlation, the greater the MZ concordance and vice versa; MZ correlations less than 1.0 indicate variables are appropriate for use in NSE analyses. For the HV5 sample (Table 4.4), MZ twin correlations ranged from .18 for Productive language to .82 for Formal Language; for environmental measures, MZ twin correlations ranged from .49 for parent-observed positivity (i.e., PARCHISY) to 1.0 for breastfeeding duration. For the HV7 sample (Table 4.5), MZ twin correlations ranged from .22 for Productive language to .80 for Formal Language; for environmental measures, MZ twin correlations ranged from .51 for parent-observed negativity (i.e., PARCHISY) to 1.0 for breastfeeding duration. As expected based on prior work (e.g., Harlaar et al., 2016), MZ correlations were higher for Formal language than Productive language. MZ correlations were also higher for parent self-reported positivity/negativity (i.e. PFQ) and home reading exposure compared to parent-observed positivity/negativity (i.e. PARCHISY). This is expected given that parents rated their identical twins for the PFQ and HRE measures, whereas naive trained observers rated the parent-child interactions (i.e., PARCHISY measures). Given extent of breastfeeding had a perfect correlation, indicating negligible twin differences, NSE analyses were not conducted on this environmental measure.

4.3.2 Primary analyses.

4.3.2.1 MZ difference correlations. To address the key questions about potential associations between discordance in environmental measures and language outcomes (questions 1 and 2), NSE experiences were derived via difference scores in identical twins reared within a shared environment. Tables 4.6 and 4.7 display correlations between MZ differences in language outcomes (Formal and Productive at mean ages 10 and 12) and MZ differences in birthweight, parenting, and home reading exposure. Independent sample *t*-tests were employed to directly examine potential sex/gender disparities in the difference scores; no significant effects emerged at $p < .05$ (not shown); consequently, sex/gender was not utilized as a covariate within the key

[30] As a means to examine the relation between language outcomes and general cognition, correlations were conducted between language composites and the Block Design subtest of the Wechsler Adult Intelligence Scale (WAIS). This subtest is one of the five subtests that make up the performance IQ index of the WAIS and measures visual-motor skills, ability to analyze geometric patterns, and part-whole recognition skills. As expected, significant correlations emerged between the Block Design subtest and Formal language across time points ($r = 0.26 – 0.28$, $p < .001$, one-tailed), and smaller correlations emerged with Productive language (HV5: $r = 0.17$, $p = .01$, one-tailed; HV7: $r = 0.11$, $p = .07$, one-tailed).

correlational and regression analyses. MZ difference correlations by sex/gender, however, are reported in Appendix C for descriptive purposes.

Of the 24 total correlation coefficients for the full unselected samples, five reached significance and one reached marginal significance, which is greater than the one significant correlation that would be expected by chance. Effect sizes ranged from small (i.e., PFQ-Positivity discordance predicted 4% of Formal language discordance at HV7) to medium (i.e., PFQ-Negativity discordance predicted 20% of Formal language discordance at both HV5 & HV7).

4.3.2.1.1 RQ1: To what extent are MZ differences in environmental measures associated with MZ differences in language outcomes at mean age 10 (i.e., HV5)? Regarding the initial research question addressing correlations between MZ differences in environmental measures and differential language outcomes at the HV5 time point (Table 4.6), only one significant effect emerged for the full sample. Specifically, Formal language discordance significantly correlated ($r(76) = -.45$, $p < .001$) with discordance in parent self-reported (i.e., PFQ) negativity, with the twin with more parent self-reported negativity at age 7 demonstrating less advanced Formal language skills at age 10. This moderate association accounted for 20% of the variance.

4.3.2.1.2 RQ2. To what extent do associations between discordance in candidate environmental measures and language outcomes replicate across time points (i.e., HV5 and HV7)? Regarding the second research question, the same correlations across environmental measures were examined in regard to language outcomes at a later time point (HV7); see Table 4.7. Specifically, the moderate negative association between discordance in parent self-reported (i.e., PFQ) negativity and Formal language was replicated at this later time point, with the same 20% effect size ($r(71) = -.45$, $p < .001$). Of interest, three additional significant and one marginally significant correlations emerged in the full sample at this later time point that were not observed at HV5. Specifically, two correlations reached significance for Formal language discordance: discordance in birthweight ($r(103) = .23$, $p = .008$) and Home Reading Exposure at mean age 7 ($r(71) = .21$, $p = .04$), with small effect sizes of 5% and 4% respectively. These effects suggested that the lighter born twins were more likely to have less advanced Formal language skills than their heavier-born co-twins, and the twins who received more parent-reported home reading exposure at mean age 7 displayed higher Formal language scores at mean age 12. One additional association reached marginal significance with Formal language discordance, parent self-reported (i.e., PFQ)

Table 4.6. *Correlations between MZ differences in language at mean age 10 (HV5) and MZ differences in environment.*

Environmental Measures	Formal Language	Productive Language
Birthweight	- 0.02	0.02
Parenting at 7		
PFQ-Positivity	0.13	0.02
PFQ-Negativity	**- 0.45****	- 0.09
PARCHISY-Positivity	0.03	0.01
PARCHISY-Negativity	0.05	- 0.02
Home Reading Exposure at 7	0.12	- 0.15

Note. One-tailed *p*-values: **p*<.05; ***p*<.01. *N* = 61-113 pairs. PFQ = Parent Feelings Questionnaire; PARCHISY = Parent-Child Interaction System.

Table 4.7. *Correlations between MZ differences in language at mean age 12 (HV7) and MZ differences in environment.*

Environmental Measures	Formal Language	Productive Language
Birthweight	**0.23****	**0.22***
Parenting at 7		
PFQ-Positivity	**0.19+**	- 0.08
PFQ-Negativity	**- 0.45****	0.01
PARCHISY-Positivity	- 0.07	- 0.04
PARCHISY-Negativity	0.18	0.04
Home Reading Exposure at 7	**0.21***	0.01

Note. One-tailed p-values: +*p*=.05; **p*<.05; ***p*<.01. All: *N* = 57-105 pairs. PFQ = Parent Feelings Questionnaire; PARCHISY = Parent-Child Interaction System.

differential positivity ($r(71) = .19$, $p = .05$), accounting for 4% of the variance. This relation suggested that higher parent self-reported positivity was associated with higher Formal language outcomes. In turn, only one significant correlation emerged for Productive language discordance, and that was birthweight discordance ($r(84) = .22$, $p = .02$), with a similar effect to that observed for Formal language discordance. Specifically, the lighter born twins had lower Productive language scores than their heavier-born co-twins, with birthweight predicting 5% of the variance.

In sum, HV7 revealed more significant correlations, with five significant or marginally significant correlations for the full sample at HV7, compared to only one reaching significance at HV5. Four of the five significant correlations at HV7 involved Formal language outcomes, with only one involving Productive language (i.e., birthweight discordance). The only significant association across both Formal and Productive language discordance was observed with birthweight discordance at HV7. The only significant correlation that replicated across both time points was the moderate negative relation between parent self-reported negativity and Formal language.

4.3.2.2 RQ3: To what extent does the association of discordance in environmental measures and language outcome measures differ based on the extent of language discordance? Quantile regression provides the opportunity to explore the relationship between environmental discordance and language discordance conditional on quantiles across the language discordance distribution. By examining the relationship at different quantiles, we were able to assess whether the relationship between a given NSE candidate and language differs based on the extent of language discordance. For example, birthweight discordance may not account for minor discrepancies in language outcomes, but it may account significantly for large discrepancies between twins in language outcomes. In addition to examining such nuances, a benefit of this technique is that it uses the full data set to estimate parameters.

Consistent with use of the methodology (Cade & Noon, 2003; Koenker, 2005; Koenker & Hallock, 2001; Petscher & Logan, 2014), two forms of evidence were considered to assess NSE effects at the extremes of the language distributions: a) significance testing for quantiles of interest (0.2, 0.8) as presented in Tables 4.8 and 4.9, and b) quantile slope plots for 9 selected quantiles (0.1, 0.2, …, 0.9) as presented in Figures 4.2 and 4.3. The 20^{th} and 80^{th} quantiles were selected as they represent the extremes of the distribution while increasing the number of twins represented by the selected quantiles, and decreasing the sampling variation compared with the 10^{th} and 90^{th} quantiles. For example, Cade and Noon (2003) reported sampling variation and estimation uncertainty generally increases at the tails of the distribution compared to the center of the distribution (see also Chernozhukov & Umanstev, 2001). Quantile slope plots are presented to more richly characterize rates of change across the distribution, as sampling variation can quickly change from one quantile to another, particularly at the extremes (Cade & Noon, 2003). More specifically, Tables 4.8 and 4.9 report parameter estimates (95% confidence intervals) for selected quantiles and significance tests for differences between quantiles, conditional on MZ language differences at mean age 10 and 12. Confidence intervals (CIs) that do not contain zero (the null value), are considered significantly different from 0 at the .05 level (significant values are in bold). Coefficients further from 0 represent a greater effect. The relationship between the environmental measures and the language measures for conditional quantile functions is also represented in Figures 4.2 and 4.3 through quantile slope plots. Strongest effects are those significantly different from 0 at the .8 quantile and/or significantly different between the .2 and .8 quantiles (see Tables 4.8 & 4.9), and with increased rates of change (in absolute value) as represented by the quantile

slope plots (see Figures 4.2 & 4.3). As expected, wider confidence intervals were observed for the tails of the distribution, particularly for the upper quantiles. Associated intercept plots (not shown) confirmed that twins at lower quantiles of language difference have lower language differences, and twins at higher quantiles of language difference have higher language differences.

4.3.2.2.1 NSE effects at the extremes of the language discordance distribution at mean age 10. For the HV5 sample, there were no significant coefficients at the 0.2 or the 0.8 quantiles for Formal language (Table 4.8). Significance tests for differences between the 0.2 and the 0.8 quantiles revealed an unexpected significant difference for parent self-reported differential positivity (Range = 0.05 to -0.19; Coefficient difference = 0.24), and an expected significant difference for parent self-reported differential negativity (Range = -0.05 to -0.38; Coefficient difference = 0.33). The quantile slope plots (Figure 4.2, left panel) provide further support for parent-reported differential negativity as coefficients remain fairly constant for the lower third of the quantiles (-0.01 to -0.05), moderately increase (in absolute value) for the central third of the quantiles (-0.04 to -0.18), and increase (in absolute value) by a factor greater than two in the upper third of the quantiles (-0.18 to -0.39). This indicates that parent-reported differential negativity appears to become more strongly related to Formal language difference for quantiles representing greater language differences, with increased negativity associated with less advanced language skills. In contrast, the quantile slope plot does not provide support for parent-reported differential positivity. The quantile slope plots for Formal language also demonstrate a positive trend for Home Reading Exposure discordance, with small coefficients from the 10^{th} to the 40^{th} quantile that remain fairly constant (ranging between 0.07 to 0.12), and with larger coefficients from the 50^{th} to the 90^{th} quantile (ranging between 0.17 to 0.34), with CIs at two quantiles excluding zero.

For Productive language, there was one coefficient that was significantly different from 0 at the 0.2 quantile, birthweight (Table 4.8). Quantile comparison tests also revealed unexpected significant differences between the 0.2 and the 0.8 quantiles for birthweight (Range = 0.21 to -0.69; Coefficient difference = 0.48). Except for the 20^{th} quantile, the quantile slope plot for birthweight (Figure 4.2, right panel) shows CIs that include zero across the distribution and widen for the upper third quantiles demonstrating the quantile slope plot does not provide support for differing associations between the extremes of the language distribution. Finally, the quantile slope plot for parent self-reported differential positivity conditional on productive language quantile demonstrates a positive trend, with coefficients continuously increasing from the 0.3 quantile to

Table 4.8. *Parameter estimates (95% CI) for selected quantiles and significance tests for differences between quantiles, conditional on MZ language differences at mean age 10 (HV5).*

	Quantile				Quantile Comparison 0.2 vs. 0.8	
	0.2	95% CI	0.8	95% CI	*F*	*p*
Formal Language at 10						
Birthweight	0.03	-0.04, 0.18	-0.11	-0.36, 0.22	0.47	0.49
Parenting at 7						
PFQ-Positivity	0.05	-0.23, 0.09	-0.19	-0.25, 0.26	**5.75**	**0.02**
PFQ-Negativity	-0.05	-0.19, 0.05	-0.38	-0.49, 0.04	**4.42**	**0.04**
PARCHISY-Positivity	0.04	-0.06, 0.09	-0.02	-0.13, 0.15	0.81	0.37
PARCHISY-Negativity	0.05	-0.01, 0.08	-0.01	-0.13, 0.20	0.49	0.49
Home Reading Exposure at 7	0.12	-0.10, 0.40	0.34	-0.14, 0.43	1.63	0.20
Productive Language at 10						
Birthweight	**0.21**	**0.03, 0.50**	-0.69	-0.81, 0.81	**4.62**	**0.03**
Parenting at 7						
PFQ-Positivity	0.14	-0.06, 0.49	0.53	-0.40, 0.55	2.49	0.12
PFQ-Negativity	0.22	-0.43, 0.34	0.13	-0.58, 0.14	0.12	0.73
PARCHISY-Positivity	-0.01	-0.22, 0.17	-0.06	-0.21, 0.21	0.11	0.74
PARCHISY-Negativity	-0.06	-0.15, 0.08	-0.15	-0.25, 0.43	0.48	0.49
Home Reading Exposure at 7	-0.23	-0.74, 0.44	-0.07	-0.44, 0.76	0.41	0.53

Note. PFQ = Parent Feelings Questionnaire; PARCHISY = Parent-Child Interaction System.

Table 4.9. *Parameter estimates (95% CI) for selected quantiles and significance tests for differences between quantiles, conditional on MZ language differences at mean age 12 (HV7).*

	Quantile				Quantile Comparison 0.2 vs. 0.8	
	0.2	95% CI	0.8	95% CI	*F*	*p*
Formal Language at 12						
Birthweight	0.00	-0.10, 0.22	**0.41**	**0.11, 0.60**	**6.18**	**0.01**
Parenting at 7						
PFQ-Positivity	0.09	-0.08, 0.18	**0.17**	**0.03, 0.36**	0.14	0.71
PFQ-Negativity	-0.01	-0.26, 0.10	**-0.44**	**-0.58, -0.10**	**5.28**	**0.02**
PARCHISY-Positivity	-0.06	-0.09, 0.03	-0.04	-0.22, -0.01	0.05	0.83
PARCHISY-Negativity	0.03	-0.05, 0.06	0.04	-0.16, 0.13	0.01	0.93
Home Reading Exposure at 7	0.05	-0.09, 0.13	0.26	-0.18, 0.45	1.14	0.29
Productive Language at 12						
Birthweight	0.22	-0.24, 0.51	**0.71**	**0.26, 1.61**	1.07	0.30
Parenting at 7						
PFQ-Positivity	-0.04	-0.37, 0.14	0.13	-0.07, 0.43	0.48	0.49
PFQ-Negativity	0.00	-0.15, 0.53	0.33	-0.60, 0.66	1.26	0.26
PARCHISY-Positivity	-0.06	-0.13, 0.00	-0.10	-0.25, 0.37	0.04	0.84
PARCHISY-Negativity	0.00	-0.08, 0.04	0.15	-0.05, 0.25	1.47	0.23
Home Reading Exposure at 7	0.11	-0.31, 0.51	-0.24	-0.38, 0.66	0.60	0.44

Note. PFQ = Parent Feelings Questionnaire; PARCHISY = Parent-Child Interaction System

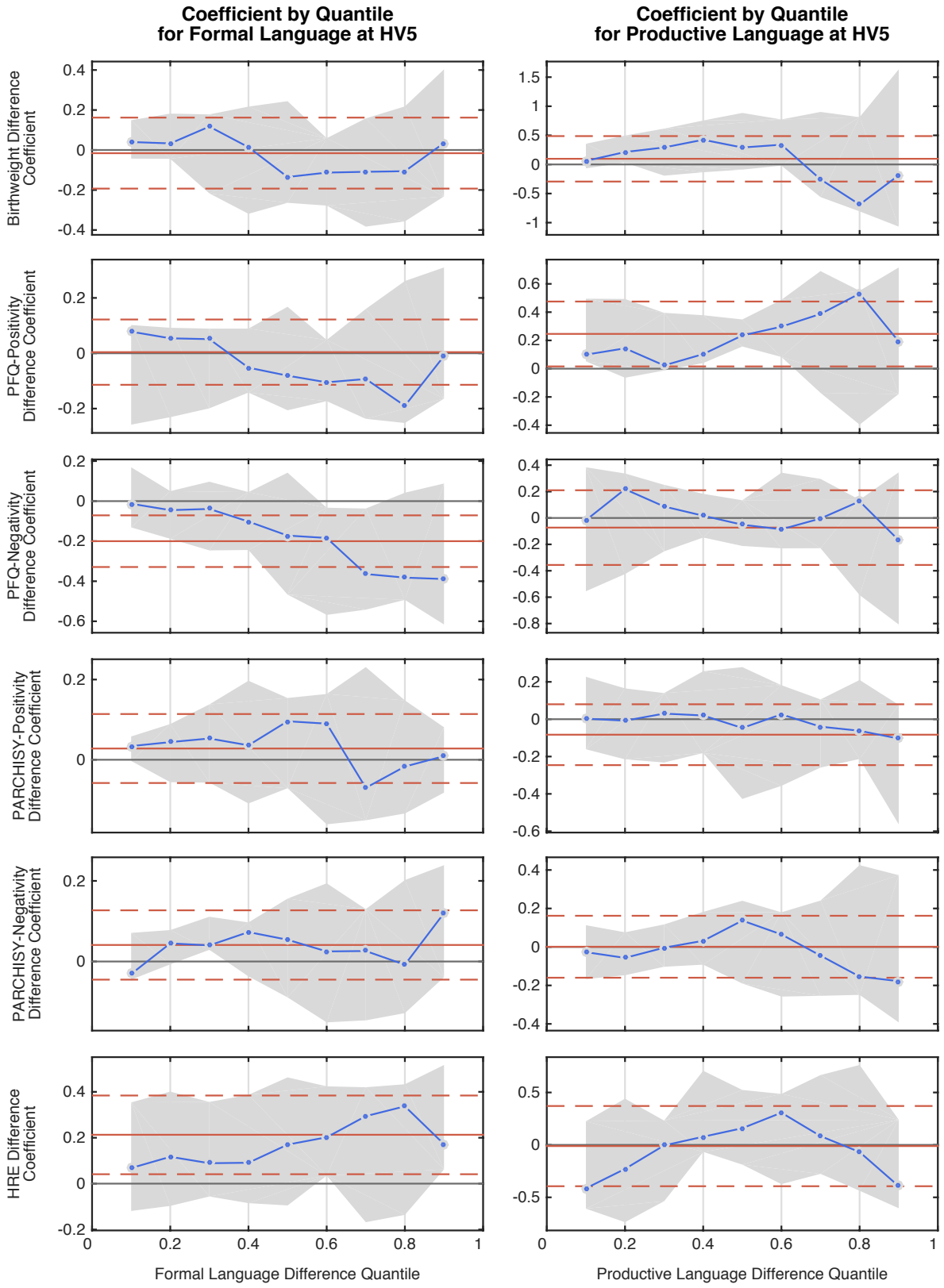


Figure 4.2. Quantile regression on language difference at mean age 10 (HV5) for environmental measures. Coefficients are represented by blue dots/line; gray shaded area indicates 95% confidence interval (CI). Ordinary least squares regression is included for comparison and represented by the red solid line; red dashed lines indicate 95% CI. PFQ = Parent Feelings Questionnaire; PARCHISY = Parent-Child Interaction System; HRE = Home Reading Exposure.

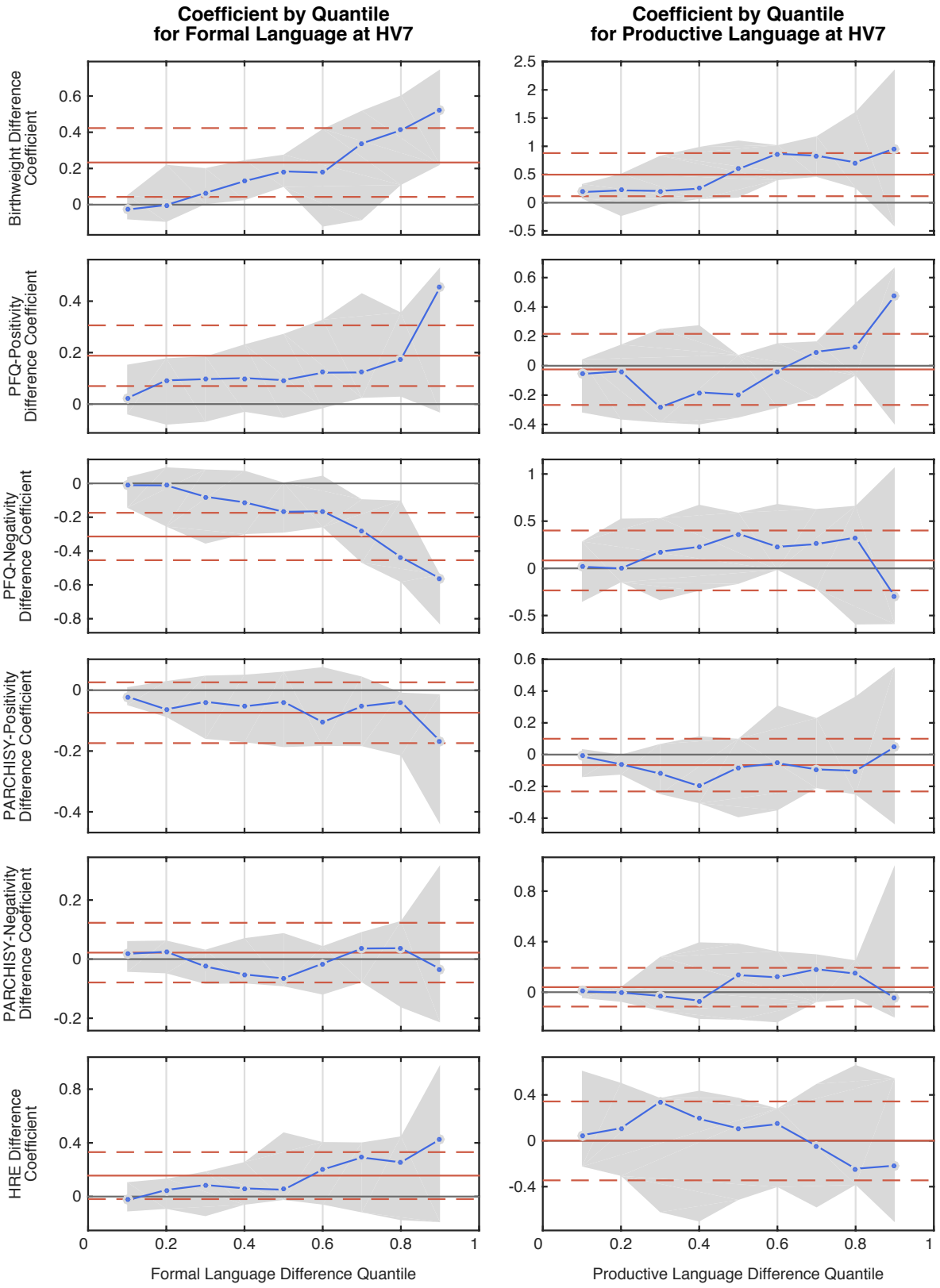


Figure 4.3. Quantile regression on language difference at mean age 12 (HV7) for environmental measures. Coefficients are represented by blue dots/line; gray shaded area indicates 95% confidence interval (CI). Ordinary least squares regression is included for comparison and represented by the red solid line; red dashed lines indicate 95% CI. PFQ = Parent Feelings Questionnaire; PARCHISY = Parent-Child Interaction System; HRE = Home Reading Exposure.

the 0.8 quantile from 0.02 to 0.53, with confidence bands excluding zero for the central third of the quantiles.

4.3.2.2.2 NSE effects at the extremes of the language discordance distribution at mean age 12. For the HV7 sample, coefficients were significantly different from zero at the 0.8 Formal language difference quantile for birthweight, PFQ-Positivity and PFQ-negativity (Table 4.9) but not at the 0.2 quantile. Significance tests for differences between the 0.2 and the 0.8 quantiles revealed significant differences for birthweight discordance (Range = 0.00 to 0.41; Coefficient difference = 0.41) and for parent self-reported differential negativity (Range = -0.01 to -0.44; Coefficient difference = -0.43) only. The quantile slope plots (Figure 4.3, left panel) provide strong support for stronger NSE effects for twin pairs more discordant in language for birthweight discordance and parent self-reported differential negativity. In both cases, coefficients remain closer to zero for the lower third of the quantiles (Bw: -0.03 to 0.07; PFQ-Neg: -0.01 to -0.08), moderately increase (in absolute value) for the central third of the quantiles (Bw: 0.07 to 0.18; PFQ-Neg: -0.08 to -0.17), and increase (in absolute value) by a factor greater than two in the upper third of the quantiles (Bw: 0.18 to 0.52; PFQ-Neg: -0.17 to -0.56). This indicates that birthweight discordance and parent self-reported differential negativity appear to become more strongly related to Formal language difference for quantiles representing greater language differences, with decreased birthweight and increased parent self-reported negativity associated with less advanced language skills. In contrast, the quantile slope plot provides moderate support for parent-reported differential positivity, with coefficients gradually increasing from 0.02 at the 10^{th} quantile to 0.17 at the 80^{th} quantile, and rapidly increasing to 0.45 at the 90^{th} quantile, with CIs at two upper quantiles excluding zero. This indicates parent-reported differential positivity becomes more strongly associated with Formal language difference for the upper quantiles, with increased parent self-reported positivity related to more advanced language skills. Although the quantile slope plots for Formal language also demonstrate a slight positive trend for home reading exposure discordance, all CIs include zero (i.e., no significant differences emerged across the distribution).

For Productive language, there was one coefficient that was significantly different from 0 at the 0.8 quantile, birthweight (Table 4.9). Quantile comparison tests did not reveal any significant differences between the 0.2 and the 0.8 quantiles. The quantile slope plots (Figure 4.3, right panel) provide further support for birthweight discordance as coefficients remain fairly constant for the

lower third of the quantiles (ranging between 0.19 to 0.22), and become larger for the central and upper third of the quantiles (ranging between 0.25 to 0.96).

***

In sum, in relation to RQ3, strongest support for increased NSE effects at the extreme of the language discordance distribution was found for measures that demonstrated a) significant differences from 0 at the .8 quantile and/or significant differences between the .2 and .8 quantiles, and b) increased rates of change (in absolute value) as represented by the quantile slope plots. At mean age 10 (HV5), this pattern of effects was seen for parent self-reported differential negativity conditional on Formal language difference quantiles. At mean age 12 (HV7) this same pattern of effects was observed for parent self-reported differential negativity, with additional effects observed for discordance in parent self-reported positivity and birthweight, all conditional on Formal language difference quantiles. In addition, NSE effects were also observed for birthweight discordance conditional on Productive language difference quantiles at this later home visit. As such, only effects for parent self-reported differential negativity conditional on Formal language difference replicated across home visits. Other effects worth highlighting include those with increased rates of change (in absolute value), with confidence intervals excluding zero in at least two central to upper quantiles across the distribution. This pattern of effects was seen at mean age 10 only (HV5), for HRE discordance conditional on Formal language difference quantiles, and parent self-reported differential positivity conditional on Productive language difference quantiles.

4.4. Discussion

This study is the first to use an MZ difference approach to examine candidate nonshared environmental effects on language development. Overall, distinct patterns of effects were observed based on form of assessment, child age, and extent of MZ pair discordance. Of the environmental variables considered, twin differences in parent self-reported negativity demonstrated the most robust effect, accounting for 20% of variance in twin Formal language discordance at both time points (at mean ages 10 and 12 years). Significant but smaller effects (4-5%) also emerged for twin discordance in birthweight and home reading exposure, with marginal significance for parent self-reported positivity in association with language outcomes, primarily Formal language, at the later time point (HV7). In general, quantile regression findings indicated stronger NSE effects as twin

language discordance increased. Stronger NSE effects with increasing language discordance is consistent with prior MZ difference studies (e.g., Asbury, Dunn, Pike, & Plomin, 2003; Asbury et al., 2006a). The overall modest effect sizes are consistent with other studies using the MZ difference approach (e.g., Asbury et al., 2006a; Mullineaux et al., 2009). Similar to candidate genes, candidate nonshared environmental influences may have small effect sizes in isolation (e.g., Plomin, DeFries, Craig, & McGuffin, 2003) but still offer viable pathways for intervention and community support. The remaining discussion focuses on a) differential effects across data sources, b) study limitations, and c) implications for future research.

4.4.1 Differential effects across data sources.

4.4.1.1 Differential effects based on form of assessment.

4.4.1.1.1 Differences across language measures. Consistent with DeThorne and colleagues (2008) and Harlaar and colleagues (2016), NSE influences differed in degree and significance based on form of language assessment (Formal vs. Productive). Based on prior work from WRRMP which reported higher nonshared environmental effects on Productive rather than Formal language (i.e., Harlaar et al., 2016), we anticipated stronger influences of the NSE candidates on Productive language. Although our pattern of MZ correlations suggested greater overall NSE effect for Productive relative to Formal language (comparable to findings from Harlaar et al., 2016), the effects of our candidate nonshared environmental measures were primarily observed for Formal language. More specifically, only birthweight demonstrated significant associations with both forms of language measurement and those effects only reached significance at the later time point (mean age 12). At least three interpretations for the differing results across Formal and Productive language measures have been previously postulated, none of which are mutually exclusive (e.g., Harlaar et al., 2016; Mahurin-Smith, DeThorne, & Petrill, in press). First, there are modality differences between the two composites; Formal language includes both receptive and expressive language measures, while Productive language is considered primarily expressive in nature (although elements of the narrative tasks also draw to some extent on language comprehension). Second, the nature of the tasks differ in terms of adult-directedness and flexibility. Consistent with standardized test procedures, Formal language tasks are more adult-directed with specific items designed explicitly to differentiate children's abilities. In contrast, the language sample measures, even though elicited through stories from standardized tools (TNL,

TLC-E) allow for more linguistic flexibility/creativity. For example, if a child can't recall a specific vocabulary word while telling a story, they can simply choose a synonym without "penalty". Third, the two forms of assessment likely differ in regard to the extent they draw upon nonlinguistic skills/factors such as executive functioning (e.g., see Mahurin-Smith et al., in press for attention mediating effects on Formal language), motivation, compliance, stress, and sociocultural practices (e.g., Allan, 1992; Fleege, Charlesworth, Burts, & Hart, 1992; Nippold, 2009; Speltz, DeKlyen, Calderon, Greenberg, & Fisher, 1999).

4.4.1.1.2 Differences across parenting measures. Self-reported parenting findings suggest decreasing negativity and enhancing positivity, as defined herein, is likely to be beneficial for children, at least children from some cultural-linguistic groups. However, it is interesting to note that similar NSE effects did not emerge for *parent-observed* measures of positivity/negativity. It was particularly surprising to find robust significant associations emerged for self-reported parenting negativity (i.e., PFQ), but not for observed parenting negativity (i.e., PARCHISY), given both measures are designed to measure parenting. One possibility is that the PFQ and the PARCHISY are measuring different components of parenting; whereas the PFQ focuses more on parenting affect (e.g., Sometimes I am not happy about my relationship with [this child]), the negativity measure of PARCHISY is a measure of both parenting affect and content/control. A second possibility is that the null results were related to the fact that parent-observed negativity ratings were highly positively skewed, with average scores of 1, indicating most parents did not display any negative content/control or negative affect during the cooperative-play tasks. It is also possible that parent-child dyads may act differently in response to being videotaped/observed (i.e., observer effect), and cultural norms on middle-class White/European Americans may pressure parents into displaying less negativity, especially for short observed tasks (i.e., 20 minutes of structured interaction). A related possibility is that the results are a reflection of the cultural homogeneity of our sample; our sample did not represent a diverse set of sociocultural groups or a diverse set of parenting styles, such as neglectful parenting. Finally, it is possible that PARCHISY is not as sensitive in picking up the smaller differences detected by the PFQ.

4.4.1.2 Differential effects across time points. This study demonstrated more significant NSE associations with language at mean age 12 relative to mean age 10, with birthweight and Home Reading Exposure reaching statistical significance, and parent self-reported positivity reaching marginal significance. This could be viewed as consistent with prior studies reporting increased

NSE influences during adolescence relative to the childhood years (DeThorne et al., 2008, 2012; Harlaar et al., 2016). The reasons for such an increase are less clear, but we provide three possible interpretations, none of which are mutually exclusive. First, prior behavioral genetic work (e.g., Scarr & McCartney, 1983) has suggested that NSE effects increase generally with age as individuals gain more freedom to personalize their own environments according to their individual proclivities. Even though in the present case genetic proclivities were controlled within twin pair, genomes may be differentially expressed as a result of differences in internal and external environmental factors, which in turn can shape how individuals respond to other environmental factors. This complex web of development leads to increasingly individualized trajectories over time.

A second possible contributor to the increase in NSE effects at the later time point relates to the onset of adolescence, a sensitive period of development marked by neurological, biological, physical, social, emotional, and cognitive changes. Brain development during adolescence is associated with structural and functional maturation of language brain regions, and increased plasticity, which suggests more receptivity to environmental influences (APA, 2002; Dahl, 2004; Huttenlocher, 1994; Peper et al., 2009; Steinberg, 2005). Whereas at age 10 children are on the cusp of childhood and adolescence, 12-yr-old children are in the midst of puberty and adolescence, which may serve to augment NSE effects. It is also possible that the moderating factors (e.g., diet) that may mitigate the effect of NSE influences (e.g., birthweight) on language development change from late childhood to early adolescence.

Third and finally, it is possible that task demands on language assessments shift somewhat between time points. As an example, both 10- and 12-years olds begin at the same start point on the CELF subtests, therefore requiring the older child to complete more items in order to achieve the same standard score. Consequently, it's possible that CELF scores at the 12-year-old time point are more influenced by nonlinguistic skills/factors such as sustained attention capabilities than at the 10-year-old time point.

4.4.2 Limitations. Two key limitations are worth highlighting from the present study. First, the quasi-experimental nature of twin-difference methodology does not assess causality or the directionality of any potential associations. For example, it is likely that discordance in parent self-reported negativity is linked to bidirectional effects in which differential child attributes influence parenting and differential parenting style influences child attributes (e.g., Caspi et al., 2004;

Deater-Deckard, 2000). Likewise, the association between parenting style and child attributes could be mediated by other factors, such as school experiences. Such nonshared environmental factors can influence epigenetic processes and ultimately result in phenotypic differences. A second key limitation relates to the relative homogeneity of our sample in terms of race/ethnicity, SES, parental education, family structure, and child ability. Consequently, to the extent these variables correlate with differences in the environmental measures studied (i.e., birthweight, parenting, home reading exposure, and breastfeeding), our effect sizes may be an under- or overestimation. For example, consistent with the quantile regression findings, NSE effects may be stronger for children at increased risk of language impairments. There is also evidence that the same environmental measure may have differential effects within different demographic groups, such as is reported for negative-intrusive parenting and child language development in African American relative to European American families (Pungello et al., 2009).

4.4.3 Implications for Future Research. Despite limitations, the present study offers a unique advantage compared to many studies of environmental effects on child language development: it specifically controls for sex/gender, age, and genetic effects. Within this context, parenting, birthweight, and home reading exposure all emerged as significant nonshared environmental influences on language development. Future work should continue to study the path of individualized developmental trajectories and attempt to replicate effects using a variety of methodologies (e.g., Asbury, Dunn & Plomin, 2006), including use of other MZ datasets with increased power, varied ages, different forms of assessment, and more demographically diverse families. Moreover, additional candidate NSE influences should be examined; in particular, nutrition, stress, illness, accidents, and exposure to toxins are all promising candidates based on prior work (e.g., Asbury et al., 2006b; Rogers et al., 2015; Verduci et al., 2014). Although our breastfeeding measure did not demonstrate adequate discordance for NSE analyses, further research examining breastfeeding specifically is warranted. Finally, to the extent that the effects for the environmental measures within the present study are causal, further research should work to elucidate the underlying mechanisms. For example, genetically sensitive intervention studies can shed more light into the causal relationship between candidate NSE influences such as parenting style and language outcomes. Studies exploring associations between differential experiences and behaviors and epigenetic differences are also needed (e.g., see Wong et al., 2014). Assessment of both mediating factors that may explain a given relationship, and moderating

factors that may mitigate the impact of an environmental experience on brain development will also contribute to the characterization of NSE effects on language development. Findings from the present study contribute to the broader literature on the etiology of child language differences by highlighting the complex web of gene and environmental contributions to development over time, and offering evidence of potential nonshared environmental factors worthy of further examination.

Acknowledgments: This project was funded by National Institute of Child Health and Human Development grants HD38075, HD46167, and HD050307 (PI: Stephen Petrill). I am appreciative of all participating families, and affiliated research staff who assisted with data collection and data management, specifically WRRMP co-investigators & associated teams including Laura DeThorne and the Child Language & Literacy Lab, Steve Petrill and the Learning Disabilities Innovation Hub, and Kirby Deater-Deckard and the Gene-Environment Processes Lab. Finally, I am thankful to Enric Xargay for thoughtful discussions.

References

Abel, K. M., Dalman, C., Svensson, A. C., Susser, E., Dal, H., Idring, S., ... Magnusson, C. (2013). Deviance in fetal growth and risk of autism spectrum disorder. *American Journal of Psychiatry*, *170*(4), 391-398. http://dx.doi.org/10.1176/appi.ajp.2012.12040543

Adegboye, A. R. A., & Heitmann, B. L. (2008). Accuracy and correlates of maternal recall of birthweight and gestational age. *BJOG: An International Journal of Obstetrics & Gynaecology*, *115*(7), 886-893. doi:10.1111/j.1471-0528.2008.01717.x

Agostin, T. M., & Bain, S. K. (1997). Predicting early school success with developmental and social skills screeners. *Psychology in the Schools*, *34*(3), 219-228. doi: 10.1002/(SICI)1520-6807(199707)34:3<219::AID-PITS4>3.0.CO;2-J

Al-Farsi, Y. M., Al-Sharbati, M. M., Waly, M. I., Al-Farsi, O. A., Al-Shafaee, M. A., Al-Khaduri, M. M., ... Deth, R. C. (2012). Effect of suboptimal breast-feeding on occurrence of autism: a case–control study. *Nutrition*, *28*(7), e27-e32. https://doi.org/10.1016/j.nut.2012.01.007

Allan, A. (1992). Development and validation of a scale to measure test-wiseness in EFL/ESL reading test takers. *Language Testing, 9*(2), 101–119. https://doi.org/10.1177/026553229200900201

Amato, P. R., & Fowler, F. (2002). Parenting practices, child adjustment, and family diversity. *Journal of marriage and family*, *64*(3), 703-716. 10.1111/j.1741-3737.2002.00703.x

American Academy of Pediatrics Committee on Environmental Health. (1997). Noise: a hazard for the fetus and newborn. *Pediatrics, 100*(4), 724–727. doi: 10.1542/peds.100.4.724

American Academy of Pediatrics. (2012). Breastfeeding and the use of human milk [Policy statement]. *Pediatrics, 129* (3), e827–e841. doi: 10.1542/peds.2011-3552

American Psychological Association. (2002). Developing adolescents: A reference for professionals. Washington, DC, USA: American Psychological Association.

Ananth, C. V., Demissie, K., & Hanley, M. L. (2003). Birth weight discordancy and adverse perinatal outcomes among twin gestations in the United States: the effect of placental abruption. *American journal of obstetrics and gynecology*, *188*(4), 954-960. http://dx.doi.org/10.1067/mob.2003.210

Anderson, J. W., Johnstone, B. M., & Remley, D. T. (1999). Breast-feeding and cognitive development: a meta-analysis. *The American journal of clinical nutrition*, *70*(4), 525-535.

Armson, B. A., O'Connell, C., Persad, V., Joseph, K. S., Young, D. C., & Baskett, T. F. (2006). Determinants of perinatal mortality and serious neonatal morbidity in the second twin. *Obstetrics & Gynecology*, *108*(3, Part 1), 556-564. doi: 10.1097/01.AOG.0000227747.37184.0a

Asbury, K., Dunn, J. F., & Plomin, R. (2006a). Birthweight-discordance and differences in early parenting relate to monozygotic twin differences in behaviour problems and academic achievement at age 7. *Developmental science*, *9*(2), F22-F31. Doi: 10.1111/j.1467-7687.2006.00469.x

Asbury, K., Dunn, J. F., & Plomin, R. (2006b). The use of discordant MZ Twins to generate hypotheses regarding non-shared environmental influence on anxiety in middle childhood. *Social Development*, *15*(3), 564-570. doi:10.1111/j.1467-9507.2006.00356.x

Asbury, K., Dunn, J. F., Pike, A., & Plomin, R. (2003), Nonshared environmental influences on individual differences in early behavioral development: A monozygotic twin differences study. *Child Development, 74*(3), 933-943. doi:10.1111/1467-8624.00577

Aylward, G. P. (2002). Cognitive and neuropsychological outcomes: more than IQ scores. *Mental retardation and developmental disabilities research reviews*, *8*(4), 234-240. doi:10.1002/mrdd.10043

Bahl, R., Frost, C., Kirkwood, B. R., Edmond, K., Martines, J., Bhandari, N., & Arthur, P. (2005). Infant feeding patterns and risks of death and hospitalization in the first half of infancy: multicentre cohort study. *Bulletin of the World Health Organization*, *83*(6), 418-426. http://dx.doi.org/10.1590/S0042-96862005000600009

Baldwin, A. L., Baldwin, C., & Cole, R. E. (1990). Stress resistant families and stress resistant children. In J. E. Roll, A. S. Masten, D. Cicchetti, K. H. Nuechterlein, & S. Weintraub (Eds.), *Risk and protective factors in the development of psychopathology* (pp. 257–280). New York, NY, USA: Cambridge University Press.

Baumwell, L., Tamis-LeMonda, C. S., & Bornstein, M. H. (1997). Maternal verbal sensitivity and child language comprehension. *Infant Behavior and Development*, *20*(2), 247-258. https://doi.org/10.1016/S0163-6383(97)90026-6

Belfort, M. B., Anderson, P. J., Nowak, V. A., Lee, K. J., Molesworth, C., Thompson, D. K., … Inder, T. E. (2016). Breast milk feeding, brain development, and neurocognitive outcomes: a 7-year longitudinal study in infants born at less than 30 weeks' gestation. *The Journal of Pediatrics*, *177*, 133-139. https://doi.org/10.1016/j.jpeds.2016.06.045

Ben-Dat Fisher, D., Serbin, L. A., Stack, D. M., Ruttle, P. L., Ledingham, J. E., & Schwartzman, A. E. (2007). Intergenerational predictors of diurnal cortisol secretion in early childhood. *Infant and Child Development*, *16*(2), 151-170. doi: 10.1002/icd.474

Boulet, S. L., Alexander, G. R., Salihu, H. M., & Pass, M. (2003). Macrosomic births in the United States: determinants, outcomes, and proposed grades of risk. *American journal of obstetrics and gynecology*, *188*(5), 1372-1378. http://dx.doi.org/10.1067/mob.2003.302

Bradley, R. H., & Corwyn, R. F. (2000). Moderating effect of perceived amount of family conflict on the relation between home environmental processes and the well-being of adolescents. *Journal of Family Psychology*, *14*(3), 349-364.

Britton, J. R., Britton, H. L., & Gronwaldt, V. (2006). Breastfeeding, sensitivity, and attachment. *Pediatrics*, *118*(5), e1436-e1443. doi: 10.1542/peds.2005-2916

Bu'Lock, F., Woolridge, M. W., & Baum, J. D. (1990). Development of co-ordination of sucking, swallowing and breathing: Ultrasound study of term and preterm infants. *Developmental Medicine & Child Neurology*, *32*(8), 669-678. doi:10.1111/j.1469-8749.1990.tb08427.x

Burgess, S. R., Hecht, S. A., & Lonigan, C. J. (2002). Relations of the home literacy environment (HLE) to the development of reading-related abilities: A one-year longitudinal study. *Reading Research Quarterly*, *37*(4), 408-426. doi: 10.1598/RRQ.37.4.4

Burke, V., Beilin, L. J., Simmer, K., Oddy, W. H., Blake, K. V., Doherty, D., ... Stanley, F. J. (2005). Breastfeeding and overweight: Longitudinal analysis in an Australian birth cohort. *The Journal of pediatrics*, *147*(1), 56-61. https://doi.org/10.1016/j.jpeds.2005.03.038

Cade, B. S., & Noon, B. R. (2003). A gentle introduction to quantile regression for ecologists. *Frontiers in Ecology and the Environment*, *1*(8), 412-420. doi: 10.1890/1540-9295(2003)001[0412:AGITQR]2.0.CO;2

Carroll, J. E., Gruenewald, T. L., Taylor, S. E., Janicki-Deverts, D., Matthews, K. A., & Seeman, T. E. (2013). Childhood abuse, parental warmth, and adult multisystem biological risk in the Coronary Artery Risk Development in Young Adults study. *Proceedings of the National Academy of Sciences*, *110*(42), 17149-17153. doi: 10.1073/pnas.1315458110

Caspi, A., Moffitt, T. E., Morgan, J., Rutter, M., Taylor, A., Arseneault, L., ... Polo-Tomas, M. (2004). Maternal expressed emotion predicts children's antisocial behavior problems: using monozygotic-twin differences to identify environmental effects on behavioral development. *Developmental psychology*, *40*(2), 149. http://psycnet.apa.org/doi/10.1037/0012-1649.40.2.149

Catts, H. W., Fey, M. E., Zhang, X., & Tomblin, J. B. (1999). Language basis of reading and reading disabilities: Evidence from a longitudinal investigation. *Scientific studies of reading*, *3*(4), 331-361. http://dx.doi.org/10.1207/s1532799xssr0304_2

Chall, J. S. (1983). *Stages of reading development*. New York, NY, USA: McGraw-Hill.

Chernozhukov, V., & Umantsev, L. (2001). Conditional value-at-risk: Aspects of modeling and estimation. *Empirical Economics*, *26*(1), 271-292.

Colen, C. G., & Ramey, D. M. (2014). Is breast truly best? Estimating the effects of breastfeeding on long-term child health and wellbeing in the United States using sibling comparisons. *Social Science & Medicine*, *109*, 55-65. https://doi.org/10.1016/j.socscimed.2014.01.027

Connell, C. M., & Prinz, R. J. (2002). The impact of childcare and parent–child interactions on school readiness and social skills development for low-income African American children. *Journal of School Psychology*, *40*(2), 177-193. http://dx.doi.org/10.1016/S0022-4405(02)00090-0

Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for specific language impairment (SLI). *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *42*(6), 741-748. https://doi.org/10.1017/S0021963001007600

Cromley, J. G., & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, *99*(2), 311. http://dx.doi.org/10.1037/0022-0663.99.2.311

Cunningham, A. E. (2005). Vocabulary growth through independent reading and reading aloud to children. In E.H. Hiebert, M.L. Kamil, (Eds.), *Teaching and learning vocabulary: Bringing research to practice*, 45-68. Mahwah, NJ, USA: Lawrence Erlbaum Associates.

Dahl, R. E. (2004). Adolescent brain development: A period of vulnerabilities and opportunities. Keynote address. *Annals of the New York Academy of Sciences, 1021*(1), 1–22. doi:10.1196/annals.1308.001

Dallaire, D. H., & Weinraub, M. (2005). The stability of parenting behaviors over the first 6 years of life. *Early Childhood Research Quarterly*, *20*(2), 201-219.

Daniels, M. C., & Adair, L. S. (2005). Breast-feeding influences cognitive development in Filipino children. *The Journal of Nutrition*, *135*(11), 2589-2595.

Darling, N., & Steinberg, L. (1993). Parenting style as context: An integrative model. *Psychological bulletin*, *113*(3), 487. http://dx.doi.org/10.1037/0033-2909.113.3.487

De Bellis, M. D. (2001). Developmental traumatology: The psychobiological development of maltreated children and its implications for research, treatment, and policy. *Development and psychopathology*, *13*(03), 539-564.

de Kieviet, J. F., Zoetebier, L., Van Elburg, R. M., Vermeulen, R. J., & Oosterlaan, J. (2012). Brain development of very preterm and very low-birthweight children in childhood and adolescence: A meta-analysis. *Developmental Medicine & Child Neurology*, *54*(4), 313-323. doi:10.1111/j.1469-8749.2011.04216.x

Deater-Deckard, K. (1996). The Parent Feelings Questionnaire. *London, UK: Institute of Psychiatry*.

Deater-Deckard, K. (2000). Parenting and child behavioral adjustment in early childhood: A quantitative genetic approach to studying family processes. *Child Development, 71*(2), 468-484. doi: 10.1111/1467-8624.00158

Deater-Deckard, K., Pylas, M. V., & Petrill, S. A. (1997). Parent-Child Interaction System (PARCHISY). *London, UK: Institute of Psychiatry.*

Deater-Deckard, K., Pike, A., Petrill, S. A., Cutting, A. L., Hughes, C., & O'Connor, T. G. (2001). Nonshared environmental processes in social-emotional development: an observational study of identical twin differences in the preschool period. *Developmental Science*, *4*(2), F1-F6. doi: 10.1111/1467-7687.00157

Dee, D. L., Li, R., Lee, L. C., & Grummer-Strawn, L. M. (2007). Associations between breastfeeding practices and young children's language and motor skill development. *Pediatrics*, *119*(Supplement 1), S92-S98. doi: 10.1542/peds.2006-2089N

Deoni, S. C., Dean, D. C., Piryatinsky, I., O'Muircheartaigh, J., Waskiewicz, N., Lehman, K., ... Dirks, H. (2013). Breastfeeding and early white matter development: a cross-sectional study. *Neuroimage*, *82*, 77-86. https://doi.org/10.1016/j.neuroimage.2013.05.090

Der, G., Batty, G. D., & Deary, I. J. (2006). Effect of breast feeding on intelligence in children: prospective study, sibling pairs analysis, and meta-analysis. *Bmj*, *333*(7575), 945. https://doi.org/10.1136/bmj.38978.699583.55

DeThorne, L. S., Deater-Deckard, K., Mahurin-Smith, J., Coletto, M. K., & Petrill, S. A. (2011). Volubility as a mediator in the associations between conversational language measures and child temperament. *International Journal of Language & Communication Disorders*, *46*(6), 700-713. doi:10.1111/j.1460-6984.2011.00034.x

DeThorne, L. S., Harlaar, N., Petrill, S. A., & Deater-Deckard, K. (2012). Longitudinal stability in genetic effects on children's conversational language productivity. *Journal of Speech, Language, and Hearing Research*, *55*(3), 739-753. doi:10.1044/1092-4388(2011/11-0014)

DeThorne, L. S., Hart, S. A., Petrill, S. A., Deater-Deckard, K., Thompson, L. A., Schatschneider, C., & Davison, M. D. (2006). Children's history of speech-language difficulties: Genetic influences and associations with reading-related measures. *Journal of Speech, Language, and Hearing Research*, *49*(6), 1280-1293. doi:10.1044/1092-4388(2006/092)

DeThorne, L. S., Petrill, S. A., Hart, S. A., Channell, R. W., Campbell, R. J., Deater-Deckard, K., ... & Vandenbergh, D. J. (2008). Genetic effects on children's conversational language use. *Journal of Speech, Language, and Hearing Research*, *51*(2), 423-435. doi:10.1044/1092-4388(2008/031)

Dibble, E., & Cohen, D. J. (1974). Companion instruments for measuring children's competence and parental style. *Archives of General Psychiatry*, *30*(6), 805-815. doi:10.1001/archpsyc.1974.01760120061010

Dionne, G., Boivin, M., Séguin, J. R., Pérusse, D., & Tremblay, R. E. (2008). Gestational diabetes hinders language development in offspring. *Pediatrics*, *122*(5), e1073-e1079. doi:10.1542/peds.2007-3028

Dodds, L., Fell, D. B., Shea, S., Armson, B. A., Allen, A. C., & Bryson, S. (2011). The role of prenatal, obstetric and neonatal factors in the development of autism. *Journal of autism and developmental disorders*, *41*(7), 891-902. https://doi.org/10.1007/s10803-010-1114-8

Dodici, B. J., Draper, D. C., & Peterson, C. A. (2003). Early parent-child interactions and early literacy development. *Topics in Early Childhood Special Education*, *23*(3), 124-136. https://doi.org/10.1177/02711214030230030301

Dotterer, A. M., Iruka, I. U., & Pungello, E. (2012). Parenting, race, and socioeconomic status: Links to school readiness. *Family Relations*, *61*(4), 657-670. doi: 10.1111/j.1741-3729.2012.00716.x

Duncan, B., Ey, J., Holberg, C. J., Wright, A. L., Martinez, F. D., & Taussig, L. M. (1993). Exclusive breast-feeding for at least 4 months protects against otitis media. *Pediatrics*, *91*(5), 867-872.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., ... Sexton, H. (2007). School readiness and later achievement. *Developmental psychology*, *43*(6), 1428. http://dx.doi.org/10.1037/0012-1649.43.6.1428

Eaton-Evans, J., & Dugdale, A. E. (1986). Recall by mothers of the birth weights and feeding of their children. *Human nutrition. Applied nutrition*, *40*(3), 171-175.

Eikenes, L., Løhaugen, G. C., Brubakk, A. M., Skranes, J., & Håberg, A. K. (2011). Young adults born preterm with very low birth weight demonstrate widespread white matter alterations on brain DTI. *Neuroimage*, *54*(3), 1774-1785. http://dx.doi.org/10.1016/j.neuroimage.2010.10.037

Elbro, C., & Scarborough, H. S. (2004). Early identification. In P. Bryant & T. Nunes (Eds.), *Handbook of children's literacy*, 339-359. Dordrecht, Netherlands: Kluwer Academic Publishers.

Eriksen, W., Sundet, J. M., & Tambs, K. (2010). Birth weight standardized to gestational age and intelligence in young adulthood: a register-based birth cohort study of male

siblings. *American journal of epidemiology*, *172*(5), 530-536. https://doi.org/10.1093/aje/kwq199

Fanaroff, A. A., Stoll, B. J., Wright, L. L., Carlo, W. A., Ehrenkranz, R. A., Stark, A. R., ... Lemons, J. A. (2007). Trends in neonatal morbidity and mortality for very low birthweight infants. *American journal of obstetrics and gynecology*, *196*(2), 147-e1. http://dx.doi.org/10.1016/j.ajog.2006.09.014

Farquharson, J., Jamieson, E. C., Logan, R. W., Cockburn, F., & Patrick, W. A. (1992). Infant cerebral cortex phospholipid fatty-acid composition and diet. *The Lancet*, *340*(8823), 810-813. https://doi.org/10.1016/0140-6736(92)92684-8

Feldman, R., Weller, A., Zagoory-Sharon, O., & Levine, A. (2007). Evidence for a neuroendocrinological foundation of human affiliation: plasma oxytocin levels across pregnancy and the postpartum period predict mother-infant bonding. *Psychological Science*, *18*(11), 965-970. https://doi.org/10.1111/j.1467-9280.2007.02010.x

Fenson, L., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., & Reznick, J. S. (2000). Short-form versions of the MacArthur communicative development inventories. *Applied Psycholinguistics*, *21*(1), 95-116.

Field, C. J. (2005). The immunological components of human milk and their effect on immune development in infants. *The Journal of Nutrition*, *135*(1), 1-4.

Fleege, P. O., Charlesworth, R., Burts, D. C., & Hart, C. H. (1992). Stress begins in kindergarten: A look at behavior during standardized testing. *Journal of Research in Childhood Education,7*(1), 20–26. doi:10.1080/02568549209594836

Forget-Dubois, N., Dionne, G., Lemelin, J. P., Pérusse, D., Tremblay, R. E., & Boivin, M. (2009). Early child language mediates the relation between home environment and school readiness. *Child development*, *80*(3), 736-749. doi:10.1111/j.1467-8624.2009.01294.x

Forsythe, S. J. (2005). Enterobacter sakazakii and other bacteria in powdered infant milk formula. *Maternal & child nutrition*, *1*(1), 44-50. doi: 10.1111/j.1740-8709.2004.00008.x

Fraga, M. F., Ballestar, E., Paz, M. F., Ropero, S., Setien, F., Ballestar, M. L., ... & Boix-Chornet, M. (2005). Epigenetic differences arise during the lifetime of monozygotic twins. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(30), 10604-10609. doi: 10.1073/pnas.0500398102

Fraser, A. M., Brockert, J. E., & Ward, R. H. (1995). Association of young maternal age with adverse reproductive outcomes. *New England journal of medicine*, *332*(17), 1113-1118. doi:10.1056/NEJM199504273321701

Galton Bachrach, V. R., Schwarz, E., & Bachrach, L. R. (2003). Breastfeeding and the risk of hospitalization for respiratory disease in infancy: a meta-analysis. *Archives of pediatrics & adolescent medicine*, *157*(3), 237-243. doi:10.1001/archpedi.157.3.237

Garcia Coll, C. T. (1990). Developmental outcome of minority infants: A process-oriented look into our beginnings. *Child Development*, *61*(2), 270 – 289. doi: 10.1111/j.1467-8624.1990.tb02779.x

Gillam, R. B., & Pearson, N. A. (2004). *Test of Narrative Language: Examiner's manual.* Austin, TX, USA: Pro-Ed.

Goldsmith, H. H. (1991). A zygosity questionnaire for young twins: A research note. *Behavior genetics*, *21*(3), 257-269. doi:10.1007/BF01065819

Granger, D. A. (1998). Children's salivary cortisol, internalising behaviour problems, and family environment: Results from the Concordia Longitudinal Risk Project. *International Journal of Behavioral Development*, *22*(4), 707-728.

Griffin, E. A., & Morrison, F. J. (1997). The unique contribution of home literacy environment to differences in early literacy skills. *Early Child Development and Care*, *127*(1), 233-243. http://dx.doi.org/10.1080/0300443971270119

Grunau, R. E., Whitfield, M. F., & Davis, C. (2002). Pattern of learning disabilities in children with extremely low birth weight and broadly average intelligence. *Archives of Pediatrics & Adolescent Medicine*, *156*(6), 615-620. doi:10.1001/archpedi.156.6.615

Grunau, R. V., Kearney, S. M., & Whitfield, M. F. (1990). Language development at 3 years in pre-term children of birth weight below 1000 g. *International Journal of Language & Communication Disorders*, *25*(2), 173-182. doi:10.3109/13682829009011972

Guastella, A. J., Einfeld, S. L., Gray, K. M., Rinehart, N. J., Tonge, B. J., Lambert, T. J., & Hickie, I. B. (2010). Intranasal oxytocin improves emotion recognition for youth with autism spectrum disorders. *Biological psychiatry*, *67*(7), 692-694. https://doi.org/10.1016/j.biopsych.2009.09.020

Guthrie, J. T., & Wigfield, A. (2000). Engagement and motivation in reading. In M. L. Kamil, P.B. Mosenthal, P.D. Pearson, & R. Barr (Eds.), *Handbook of Reading Research*, 3rd ed., 403-422. New York, NY, USA: Longman.

Hanson, L. Å., Korotkova, M., & Telemo, E. (2003). Breast-feeding, infant formulas, and the immune system. *Annals of Allergy, Asthma & Immunology*, *90*(6), 59-63. https://doi.org/10.1016/S1081-1206(10)61662-6

Harder, T., Bergmann, R., Kallischnigg, G., & Plagemann, A. (2005). Duration of breastfeeding and risk of overweight: A meta-analysis. *American journal of epidemiology*, *162*(5), 397-403. https://doi.org/10.1093/aje/kwi222

Hargrave, A. C., & Sénéchal, M. (2000). A book reading intervention with preschool children who have limited vocabularies: The benefits of regular reading and dialogic reading. *Early Childhood Research Quarterly*, *15*(1), 75-90. http://dx.doi.org/10.1016/S0885-2006(99)00038-1

Harlaar, N., DeThorne, L. S., Mahurin-Smith, J., Aparicio Betancourt, M., & Petrill, S. A. (2016). Longitudinal Effects on Early Adolescent Language: A Twin Study. *Journal of Speech, Language, and Hearing Research*, *59*(5), 1059-1073. doi:10.1044/2016_JSLHR-L-15-0257

Harlaar, N., Hayiou-Thomas, M. E., Dale, P. S., & Plomin, R. (2008). Why do preschool language abilities correlate with later reading? A twin study. *Journal of Speech, Language, and Hearing Research*, *51*(3), 688-705. doi:10.1044/1092-4388(2008/049)

Harrison, L. J., & McLeod, S. (2010). Risk and protective factors associated with speech and language impairment in a nationally representative sample of 4-to 5-year-old children.

Journal of Speech, Language, and Hearing Research, *53*(2), 508-529. doi:10.1044/1092-4388(2009/08-0086)

Hart, S. A., Petrill, S. A., DeThorne, L. S., Deater-Deckard, K., Thompson, L. A., Schatschneider, C., & Cutting, L. E. (2009). Environmental influences on the longitudinal covariance of expressive vocabulary: measuring the home literacy environment in a genetically sensitive design. *Journal of Child Psychology and Psychiatry*, *50*(8), 911-919. doi: 10.1111/j.1469-7610.2009.02074.x

Hartley, R. S., & Hitti, J. (2005). Birth order and delivery interval: analysis of twin pair perinatal outcomes. *The Journal of Maternal-Fetal & Neonatal Medicine*, *17*(6), 375-380. http://dx.doi.org/10.1080/147670505000132484

Heilmann, J., Nockerts, A., & Miller, J. F. (2010). Language sampling: Does the length of the transcript matter? *Language, Speech, and Hearing Services in Schools*, *41*(4), 393-404. doi:10.1044/0161-1461(2009/09-0023)

Hurlemann, R., Patin, A., Onur, O. A., Cohen, M. X., Baumgartner, T., Metzler, S., ... Kendrick, K. M. (2010). Oxytocin enhances amygdala-dependent, socially reinforced learning and emotional empathy in humans. *Journal of Neuroscience*, *30*(14), 4999-5007. https://doi.org/10.1523/JNEUROSCI.5538-09.2010

Hutchins, T. L., Brannick, M., Bryant, J. B., & Silliman, E. R. (2005). Methods for controlling amount of talk: Difficulties, considerations and recommendations. *First Language*, *25*(3), 347-363. https://doi.org/10.1177/0142723705056376

Huttenlocher, P. R. (1994). Synaptogenesis in human cerebral cortex. In G. Dawson & K. W. Fischer (Eds.), *Human behavior and the developing brain* (pp. 137–152). New York, NY, USA: Guilford Press.

Huttenlocher, P. R. (2002). *Neural plasticity: The effects of environment on the development of the cerebral cortex*. Cambridge, MA, USA: Harvard University Press.

Innis, S. M. (2007). Dietary (n-3) fatty acids and brain development. *The Journal of nutrition*, *137*(4), 855-859.

Isaacs, E. B., Fischl, B. R., Quinn, B. T., Chong, W. K., Gadian, D. G., & Lucas, A. (2010). Impact of breast milk on intelligence quotient, brain size, and white matter development. *Pediatric research*, *67*(4), 357-362. doi:10.1203/PDR.0b013e3181d026da

Isbell, R., Sobol, J., Lindauer, L., & Lowrance, A. (2004). The effects of storytelling and story reading on the oral language complexity and story comprehension of young children. *Early childhood education journal*, *32*(3), 157-163. doi:10.1023/B:ECEJ.0000048967.94189.a3

Jacobson, L. (2001). Education department advocates clash at NAEYC meeting. *Education Week*, 12.

Jansson-Verkasalo, E., Valkama, M., Vainionpää, L., Pääkkö, E., Ilkko, E., & Lehtihalmes, M. (2004). Language development in very low birth weight preterm children: a follow-up study. *Folia phoniatrica et logopaedica*, *56*(2), 108-119. https://doi.org/10.1159/000076062

Jennische, M., & Sedin, G. (2001). Linguistic skills at 6½ years of age in children who required neonatal intensive care in 1986–1989. *Acta Paediatrica*, *90*(2), 199-212. doi:10.1111/j.1651-2227.2001.tb00285.x

Johnson, E. O., & Breslau, N. (2000). Increased risk of learning disabilities in low birth weight boys at age 11 years. *Biological psychiatry*, *47*(6), 490-500. http://dx.doi.org/10.1016/S0006-3223(99)00223-1

Kafouri, S., Kramer, M., Leonard, G., Perron, M., Pike, B., Richer, L., ... & Paus, T. (2012). Breastfeeding and brain structure in adolescence. *International journal of epidemiology*, *42*(1), 150-159. https://doi.org/10.1093/ije/dys172

Kim, P., Feldman, R., Mayes, L. C., Eicher, V., Thompson, N., Leckman, J. F., & Swain, J. E. (2011). Breastfeeding, brain activation to own infant cry, and maternal sensitivity. *Journal of child psychology and psychiatry*, *52*(8), 907-915. doi: 10.1111/j.1469-7610.2011.02406.x

Koenker, R. (2005). *Quantile regression* (No. 38). A. Chesher & M. Jackson (Eds.). New York, NY, USA: Cambridge University Press.

Koenker, R., & Bassett Jr, G. (1978). Regression quantiles. *Econometrica: Journal of the Econometric Society, 46*(1), 33-50. doi: 10.2307/1913643

Koenker, R., & Hallock, K. (2001). Quantile regression. *The Journal of Economic Perspectives*, *15*(4), 143-156. Retrieved from http://www.jstor.org/stable/2696522

Korja, R., Latva, R., & Lehtonen, L. (2012). The effects of preterm birth on mother–infant interaction and attachment during the infant's first two years. *Acta obstetricia et gynecologica Scandinavica*, *91*(2), 164-173. doi: 10.1111/j.1600-0412.2011.01304.x

Kramer, M. S., Aboud, F., Mironova, E., Vanilovich, I., Platt, R. W., Matush, L., ... Collet, J. P. (2008). Breastfeeding and child cognitive development: new evidence from a large randomized trial. *Archives of general psychiatry*, *65*(5), 578-584. doi:10.1001/archpsyc.65.5.578

Lamborn, S. D., Dornbusch, S. M., & Steinberg, L. (1996). Ethnicity and community context as moderators of the relations between family decision making and adolescent adjustment. *Child development*, *67*(2), 283-301. doi: 10.1111/j.1467-8624.1996.tb01734.x

Landry, S. H., Smith, K. E., Swank, P. R., Assel, M. A., & Vellet, S. (2001). Does early responsive parenting have a special importance for children's development or is consistency across early childhood necessary?. *Developmental psychology*, *37*(3), 387-403.

Lepage P., Van de Perre P. (2012). The immune system of breast milk: Antimicrobial and anti-inflammatory properties. In A. Kourtis, M. Bulterys (Eds.), *Human immunodeficiency virus type 1 (HIV-1) and breastfeeding: Advances in experimental medicine and biology* (vol 743, pp. 121-137). New York, NY, USA: Springer.

Leventakou, V., Roumeliotaki, T., Koutra, K., Vassilaki, M., Mantzouranis, E., Bitsios, P., ... & Chatzi, L. (2015). Breastfeeding duration and cognitive, language and motor development at 18 months of age: Rhea mother–child cohort in Crete, Greece. *J Epidemiol Community Health*, *69*(3), 232-239. http://dx.doi.org/10.1136/jech-2013-202500

Lewis, B. A., & Thompson, L. A. (1992). A study of developmental speech and language disorders in twins. *Journal of Speech, Language, and Hearing Research*, *35*(5), 1086-1094. doi:10.1044/jshr.3505.1086

Li, R., Scanlon, K. S., & Serdula, M. K. (2005). The validity and reliability of maternal recall of breastfeeding practice. *Nutrition reviews*, *63*(4), 103-110. https://doi.org/10.1111/j.1753-4887.2005.tb00128.x

Loban, W. (1976). *Language Development: Kindergarten through Grade Twelve* (NCTE Committee on Research Report No. 18). Urbana, IL, USA: National Council of Teachers of English.

Lucas, A., Morley, R., & Cole, T. J. (1998). Randomised trial of early diet in preterm babies and later intelligence quotient. *Bmj*, *317*(7171), 1481-1487. https://doi.org/10.1136/bmj.317.7171.1481

Lucas, A., Morley, R., Cole, T. J., Lister, G., & Leeson-Payne, C. (1992). Breast milk and subsequent intelligence quotient in children born preterm. *The Lancet*, *339*(8788), 261-264. https://doi.org/10.1016/0140-6736(92)91329-7

Mahurin-Smith, J. (2015). Breastfeeding and language outcomes: A review of the literature. *Journal of communication disorders*, *57*, 29-40. https://doi.org/10.1016/j.jcomdis.2015.04.002

Mahurin-Smith, J., & Ambrose, N. G. (2013). Breastfeeding may protect against persistent stuttering. *Journal of Communication Disorders*, *46*(4), 351-360. http://dx.doi.org/10.1016/j.jcomdis.2013.06.001

Mahurin-Smith, J., DeThorne, L. S., & Petrill, S. A. (in press). Longitudinal associations across prematurity, attention, and language in school-age children. *Journal of Speech, Language, and Hearing Research.*

Mahurin-Smith, J., DeThorne, L. S., Logan, J. A., Channell, R. W., & Petrill, S. A. (2014). Impact of prematurity on language skills at school age. *Journal of Speech, Language, and Hearing Research*, *57*(3), 901-916. doi:10.1044/1092-4388(2013/12-0347)

Malin, G. L., Morris, R. K., Riley, R., Teune, M. J., & Khan, K. S. (2014). When is birthweight at term abnormally low? A systematic review and meta-analysis of the association and predictive ability of current birthweight standards for neonatal outcomes. *BJOG: An International Journal of Obstetrics & Gynaecology*, *121*(5), 515-526. doi:10.1111/1471-0528.12517

Mårild, S., Hansson, S., Jodal, U., Oden, A., & Svedberg, K. (2004). Protective effect of breastfeeding against urinary tract infection. *Acta Paediatrica*, *93*(2), 164-167. doi: 10.1111/j.1651-2227.2004.tb00699.x

Martin, J. A., Hamilton, B. E., Osterman, M. J. K., Driscoll, A. K., & Mathews, T. J. (2017). Births: Final data for 2015. *National Vital Statistics Reports*, *66*(1). Hyattsville, MD, USA: National Center for Health Statistics.

Martinussen, M., Flanders, D. W., Fischl, B., Busa, E., Løhaugen, G. C., Skranes, J., ... Dale, A. M. (2009). Segmental brain volumes and cognitive and perceptual correlates in 15-year-

old adolescents with low birth weight. *The Journal of Pediatrics*, *155*(6), 848-853. http://dx.doi.org/10.1016/j.jpeds.2009.06.015

Matthews, T. J., MacDorman, M. F., & Thoma, M. E. (2015). Infant mortality statistics from the 2013 period linked birth/infant death data set. *National vital statistics reports, 64*(9), 1-30. Hyattsville, MD, USA: Centers for Disease Control and Prevention, National Center for Health Statistics.

Mayo Clinic. (2015). *Fetal macrosomia*. Last updated April 16, 2015. Retrieved from http://www.mayoclinic.org/diseases-conditions/fetal-macrosomia/basics/definition/con-20035423

McCardle, P., Scarborough, H. S., & Catts, H. W. (2001). Predicting, explaining, and preventing children's reading difficulties. *Learning Disabilities Research & Practice*, *16*(4), 230-239. doi:10.1111/0938-8982.00023

McFadyen, M., Farquharson, J., & Cockburn, F. (2003). Maternal and umbilical cord erythrocyte omega-3 and omega-6 fatty acids and haemorheology in singleton and twin pregnancies. *Archives of Disease in Childhood-Fetal and Neonatal Edition*, *88*(2), F134-F138. http://dx.doi.org/10.1136/fn.88.2.F134

Meyer-Lindenberg, A., Domes, G., Kirsch, P., & Heinrichs, M. (2011). Oxytocin and vasopressin in the human brain: social neuropeptides for translational medicine. *Nature Reviews Neuroscience*, *12*(9), 524-538. doi:10.1038/nrn3044

Miller, J. F., Freiberg, C., Holland, M. B., & Reeves, M. A. (1992). Implementing computerized language sample analysis in the public school. *Topics in Language Disorders*, *12*(2), 69-82. https://doi.org/10.1097/00011363-199202000-00008

Miller, J. F., Iglesias, A., & Nockerts, A. (2004). *Systematic Analysis of Language Transcripts* (Research Version 8) [Computer software]. Madison, WI, USA: University of Wisconsin-Madison.

Miller, J., Long, S., McKinley, N., Thormann, S., Jones, M., & Nockerts, A. (2005). *Language sample analysis II: The Wisconsin guide (Bulletin No. 5056)*. Milwaukee, WI, USA: Wisconsin Department of Public Instruction.

Mistry, R. S., Biesanz, J. C., Taylor, L. C., Burchinal, M., & Cox, M. J. (2004). Family income and its relation to preschool children's adjustment for families in the NICHD Study of Early Child Care. *Developmental psychology*, *40*(5), 727-745.

Mortensen, E. L., Michaelsen, K. F., Sanders, S. A., & Reinisch, J. M. (2002). The association between duration of breastfeeding and adult intelligence. *Jama*, *287*(18), 2365-2371. doi:10.1001/jama.287.18.2365

Mullineaux, P. Y., Deater-Deckard, K., Petrill, S. A., & Thompson, L. A. (2009). Parenting and child behaviour problems: a longitudinal analysis of non-shared environment. *Infant and Child Development*, *18*(2), 133-148. doi: 10.1002/icd.593

Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: evidence from a longitudinal study. *Developmental psychology*, *40*(5), 665. http://dx.doi.org/10.1037/0012-1649.40.5.665

Newman, S. B. (1996). Children engaging in storybook reading: The influence of access to print resources, opportunity, and parental interaction. *Early Childhood Research Quarterly*, *11*(4), 495-513. https://doi.org/10.1016/S0885-2006(96)90019-8

Nippold, M. A. (1998). *Later Language Development: The School-age and Adolescent Years* (2nd ed.). Austin, TX, USA: Pro-Ed.

Nippold, M. A. (2009). School-age children talk about chess: Does knowledge drive syntactic complexity? *Journal of Speech, Language, and Hearing Research, 52*(4), 856–871. doi:10.1044/1092-4388(2009/08-0094)

Nyaradi, A., Oddy, W. H., Hickling, S., Li, J., & Foster, J. K. (2015). The relationship between nutrition in infancy and cognitive performance during adolescence. *Frontiers in Nutrition*, *2*, 2. doi: 10.3389/fnut.2015.00002

Oddy, W. H., Kendall, G. E., Li, J., Jacoby, P., Robinson, M., De Klerk, N. H., ... Stanley, F. J. (2010). The long-term effects of breastfeeding on child and adolescent mental health: a pregnancy cohort study followed for 14 years. *The Journal of Pediatrics*, *156*(4), 568-574. https://doi.org/10.1016/j.jpeds.2009.10.020

Oddy, W. H., Robinson, M., Kendall, G. E., Li, J., Zubrick, S. R., & Stanley, F. J. (2011). Breastfeeding and early child development: a prospective cohort study. *Acta Paediatrica*, *100*(7), 992-999. doi: 10.1111/j.1651-2227.2011.02199.x

Ohio Census. (2010). *Ohio QuickFacts*. Retrieved March 18, 2017 from https://www.census.gov/quickfacts/table/PST045215/39

Peper, J. S., Schnack, H. G., Brouwer, R. M., Van Baal, G. C. M., Pjetri, E., Székely, E., . . . Hulshoff Pol, H. E. (2009). Heritability of regional and global brain structure at the onset of puberty: A magnetic resonance imaging study in 9-year-old twin pairs. *Human Brain Mapping, 30*(7), 2184–2196. doi:10.1002/hbm.20660

Perna, R., Loughan, A. R., Le, J., & Tyson, K. (2015). Gestational diabetes: Long-term central nervous system developmental and cognitive sequelae. *Applied Neuropsychology: Child*, *4*(3), 217-220. http://dx.doi.org/10.1080/21622965.2013.874951

Peterson, B. S., Vohr, B., Staib, L. H., Cannistraci, C. J., Dolberg, A., Schneider, K. C., ... Duncan, C. C. (2000). Regional brain volume abnormalities and long-term cognitive outcome in preterm infants. *Jama*, *284*(15), 1939-1947. doi:10.1001/jama.284.15.1939

Petrill, S. A., Deater-Deckard, K., Schatschneider, C., & Davis, C. (2005). Measured environmental influences on early reading: Evidence from an adoption study. *Scientific Studies of Reading*, *9*(3), 237-259. http://dx.doi.org/10.1207/s1532799xssr0903_4

Petrill, S. A., Deater-Deckard, K., Thompson, L. A., DeThorne, L. S., & Schatschneider, C. (2006). Reading skills in early readers genetic and shared environmental influences. *Journal of Learning Disabilities*, *39*(1), 48-55. doi:10.1177/00222194060390010501

Petscher, Y., & Logan, J. A. R. (2014). Quantile regression in the study of developmental sciences. *Child development*, *85*(3), 861-881. doi:10.1111/cdev.12190

Peyre, H., Bernard, J. Y., Forhan, A., Charles, M. A., De Agostini, M., Heude, B., & Ramus, F. (2014). Predicting changes in language skills between 2 and 3 years in the EDEN mother–child cohort. *PeerJ*, *2*, e335. doi: 10.7717/peerj.335

Pharoah, P. O., Stevenson, C. J., Cooke, R. W., & Stevenson, R. C. (1994). Clinical and subclinical deficits at 8 years in a geographically defined cohort of low birthweight infants. *Archives of disease in childhood*, *70*(4), 264-270. http://dx.doi.org/10.1136/adc.70.4.264

Pianta, R. C. (1997). Adult–child relationship processes and early schooling. *Early Education and Development*, *8*(1), 11-26. http://dx.doi.org/10.1207/s15566935eed0801_2

Pianta, R. C., Nimetz, S. L., & Bennett, E. (1997). Mother-child relationships, teacher-child relationships, and school outcomes in preschool and kindergarten. *Early Childhood Research Quarterly*, *12*(3), 263-280. https://doi.org/10.1016/S0885-2006(97)90003-X

Pike, A., Reiss, D., Hetherington, E. M., & Plomin, R. (1996). Using MZ differences in the search for nonshared environmental effects. *Journal of Child Psychology and Psychiatry*, *37*(6), 695-704. doi:10.1111/j.1469-7610.1996.tb01461.x

Plomin, R. E., DeFries, J. C., Craig, I. W., & McGuffin, P. E. (Eds.) (2003). *Behavioral genetics in the postgenomic era*. Washington, DC, USA: American Psychological Association.

Plomin, R., & Kovas, Y. (2005). Generalist genes and learning disabilities. *Psychological bulletin*, *131*(4), 592. http://psycnet.apa.org/doi/10.1037/0033-2909.131.4.592

Plomin, R., Asbury, K., & Dunn, J. (2001). Why are children in the same family so different? Nonshared environment a decade later. *The Canadian Journal of Psychiatry*, *46*(3), 225-233. https://doi.org/10.1177/070674370104600302

Poulsen, P., Esteller, M., Vaag, A., & Fraga, M. F. (2007). The epigenetic basis of twin discordance in age-related diseases. *Pediatric Research*, *61*, 38R-42R. doi:10.1203/pdr.0b013e31803c7b98

Pungello, E. P., Iruka, I. U., Dotterer, A. M., Mills-Koonce, R., & Reznick, J. S. (2009). The effects of socioeconomic status, race, and parenting on language development in early childhood. *Developmental psychology*, *45*(2), 544-557. http://dx.doi.org/10.1037/a0013917

Pyles, M. K., Stolz, H. R., & Macfarlane, J. W. (1935). The accuracy of mothers' reports on birth and developmental data. *Child Development*, *6*(3), 165-176. doi:10.2307/1125860

Qi, C. H., & Kaiser, A. P. (2004). Problem behaviors of low-income children with language delays: An observation study. *Journal of Speech, Language, and Hearing Research,* 47(3), 595-609. doi: 10.1044/1092-4388(2004/046).

Quigley, M. A., Hockley, C., Carson, C., Kelly, Y., Renfrew, M. J., & Sacker, A. (2012). Breastfeeding is associated with improved child cognitive development: a population-based cohort study. *The Journal of Pediatrics*, *160*(1), 25-32. https://doi.org/10.1016/j.jpeds.2011.06.035

Rahkonen, P., Heinonen, K., Pesonen, A. K., Lano, A., Autti, T., Puosi, R., ... Räikkönen, K. (2014). Mother-child interaction is associated with neurocognitive outcome in extremely low gestational age children. *Scandinavian Journal of Psychology*, *55*(4), 311-318. doi: 10.1111/sjop.12133

Ramachandrappa, A., & Jain, L. (2009). Health issues of the late preterm infant. *Pediatric Clinics of North America*, *56*(3), 565-577. http://dx.doi.org/10.1016/j.pcl.2009.03.009

Redmond, S. M., & Rice, M. L. (1998). The socioemotional behaviors of children with SLI: Social adaptation or social deviance? *Journal of Speech, Language, and Hearing Research,* 41(3), 688-700. doi: 10.1044/jslhr.4103.688.

Rice, M. L., Redmond, S. M., & Hoffman, L. (2006). Mean length of utterance in children with specific language impairment and in younger control children shows concurrent validity and stable and parallel growth trajectories. *Journal of Speech, Language, and Hearing Research*, *49*(4), 793-808. doi:10.1044/1092-4388(2006/056)

Rogers, C. R., Nulty, K. L., Aparicio Betancourt, M., & DeThorne, L. S. (2015). Causal effects on child language development: A review of studies in communication sciences and disorders. *Journal of communication disorders*,*57*, 3-15. http://dx.doi.org/10.1016/j.jcomdis.2015.06.004

Sadauskaitė-Kuehne, V., Ludvigsson, J., Padaiga, Ž., Jašinskienė, E., & Samuelsson, U. (2004). Longer breastfeeding is an independent protective factor against development of type 1 diabetes mellitus in childhood. *Diabetes/metabolism Research and Reviews*, *20*(2), 150-157. doi: 10.1002/dmrr.425

Salem, N., Moriguchi, T., Greiner, R. S., McBride, K., Ahmad, A., Catalan, J. N., & Slotnick, B. (2001). Alterations in brain function after loss of docosahexaenoate due to dietary restriction of n-3 fatty acids. *Journal of Molecular Neuroscience*, *16*(2), 299-307. https://doi.org/10.1385/JMN:16:2-3:299

Savaskan, E., Ehrhardt, R., Schulz, A., Walter, M., & Schächinger, H. (2008). Post-learning intranasal oxytocin modulates human memory for facial identity. *Psychoneuroendocrinology*, *33*(3), 368-374. https://doi.org/10.1016/j.psyneuen.2007.12.004

Scarborough, H. (2005). Developmental relationships between language and reading: Reconciling a beautiful hypothesis with some ugly facts. In H. Catts & A. Kamhi (Eds.), *The Connections Between Language and Reading Disabilities*, 3–24. Mahwah, NJ, USA: Erlbaum.

Scarborough, H. S., & Dobrich, W. (1994). On the efficacy of reading to preschoolers. *Developmental review*, *14*(3), 245-302. https://doi.org/10.1006/drev.1994.1010

Scarr, S., & McCartney, K. (1983). How people make their own environments: A theory of genotype→ environment effects. *Child development, 54*(2), 424-435. doi:10.2307/1129703

Schlaggar, B. L., & McCandliss, B. D. (2007). Development of neural systems for reading. *Annu. Rev. Neurosci.*, *30*, 475-503. doi: 10.1146/annurev.neuro.28.061604.135645

Schultz, S. T., Klonoff-Cohen, H. S., Wingard, D. L., Akshoomoff, N. A., Macera, C. A., Ji, M., & Bacher, C. (2006). Breastfeeding, infant formula supplementation, and autistic disorder: the results of a parent survey. *International breastfeeding journal*, *1*(1), 16. https://doi.org/10.1186/1746-4358-1-16

Scott, J. A., Binns, C. W., Oddy, W. H., & Graham, K. I. (2006). Predictors of breastfeeding duration: evidence from a cohort study. *Pediatrics*, *117*(4), e646-e655. doi: 10.1542/peds.2005-1991

Semel, E., Wiig, E. H., & Secord, W. A. (2003). *Clinical Evaluation of Language Fundamentals-Fourth Edition*. San Antonio, TX, USA: The Psychological Corporation.

Sénéchal, M., Lefevre, J. A., Thomas, E. M., & Daley, K. E. (1998). Differential effects of home literacy experiences on the development of oral and written language. *Reading Research Quarterly*, *33*(1), 96-116. doi: 10.1598/RRQ.33.1.5

Sénéchal, M., Lefevre, J., Hudson, E., & Lawson, E. P. (1996). Knowledge of storybooks as a predictor of young children's vocabulary. *Journal of Educational Psychology, 88*, 520–536.

Severi, F. M., Rizzo, G., Bocchi, C., D'Antona, D., Verzuri, M. S., & Arduini, D. (2000). Intrauterine growth retardation and fetal cardiac function. *Fetal diagnosis and therapy*, *15*(1), 8-19. https://doi.org/10.1159/000020969

Siddiqui, F., & McEwan, A. (2007). Twins. *Obstetrics, Gynaecology & Reproductive Medicine*, *17*(10), 289-295. http://dx.doi.org/10.1016/j.ogrm.2007.07.011

Silfverdal, S. A., Bodin, L., & Olcén, P. (1999). Protective effect of breastfeeding: an ecologic study of Haemophilus influenzae meningitis and breastfeeding in a Swedish population. *International Journal of Epidemiology*, *28*(1), 152-156. https://doi.org/10.1093/ije/28.1.152

Speltz, M. L., DeKlyen, M., Calderon, R., Greenberg, M. T., & Fisher, P. A. (1999). Neuropsychological characteristics and test behaviors of boys with early onset conduct problems. *Journal of Abnormal Psychology, 108*(2), 315–325. doi: 10.1037/0021-843X.108.2.315

Spera, C. (2005). A review of the relationship among parenting practices, parenting styles, and adolescent school achievement. *Educational Psychology Review*, *17*(2), 125-146. doi: 10.1007/s10648-005-3950-1

Steinberg, L. (2005). Cognitive and affective development in adolescence. *Trends in Cognitive Sciences, 9*(2), 69–74. doi:10.1016/j.tics.2004.12.005

Stevenson-Hinde, J., & Shouldice, A. (1995). Maternal interactions and self-reports related to attachment classifications at 4.5 years. *Child Development*, *66*(3), 583-596. doi: 10.1111/j.1467-8624.1995.tb00891.x

Stiles, J. (2008). *The fundamentals of brain development: Integrating nature and nurture*. Cambridge, MA, USA: Harvard University Press.

Stromswold, K. (2001). The heritability of language: A review and metaanalysis of twin, adoption, and linkage studies. *Language*, 647-723. http://www.jstor.org/stable/3086845

Stromswold, K. (2006). Why aren't identical twins linguistically identical? Genetic, prenatal and postnatal factors. *Cognition*, *101*(2), 333-384. http://dx.doi.org/10.1016/j.cognition.2006.04.007

Tamis-LeMonda, C. S., Bornstein, M. H., & Baumwell, L. (2001). Maternal responsiveness and children's achievement of language milestones. *Child development*, *72*(3), 748-767. doi: 10.1111/1467-8624.00313

Tamminen, T., Verronen, P., Saarikoski, S., Göransson, A., & Tuomiranta, H. (1983). The influence of perinatal factors on breast feeding. *Acta Paediatrica*, *72*(1), 9-12. doi: 10.1111/j.1651-2227.1983.tb09655.x

Thorpe, K., Rutter, M., & Greenwood, R. (2003). Twins as a natural experiment to study the causes of mild language delay: II: Family interaction risk factors. *Journal of Child Psychology and Psychiatry*, *44*(3), 342-355. doi:10.1111/1469-7610.00126

Tomblin, J. B., Smith, E., & Zhang, X. (1997). Epidemiology of specific language impairment: Prenatal and perinatal risk factors. *Journal of communication disorders*, *30*(4), 325-344. https://doi.org/10.1016/S0021-9924(97)00015-4

Tomeo, C. A., Rich-Edwards, J. W., Michels, K. B., Berkey, C. S., Hunter, D. J., Frazier, A. L., ... Buka, S. L. (1999). Reproducibility and validity of maternal recall of pregnancy-related events. *Epidemiology*, *10*(6), 774-776.

Tomizawa, K., Iga, N., Lu, Y. F., Moriwaki, A., Matsushita, M., Li, S. T., ... Matsui, H. (2003). Oxytocin improves long-lasting spatial memory during motherhood through MAP kinase cascade. *Nature Neuroscience*, *6*(4), 384-390. doi:10.1038/nn1023

U.S. Department of Health and Human Services, Health Resources and Services Administration, Maternal and Child Health Bureau. (2014). *Child Health USA 2014*. Rockville, Maryland, USA: U.S. Department of Health and Human Services. Retrieved from https://mchb.hrsa.gov/chusa14/health-status-behaviors/infants/preterm-birth-low-birth-weight.html

United Nations Children's Fund & World Health Organization. (2004). *Low birthweight: Country, regional and global estimates*. New York, NY, USA: UNICEF.

Uvnäs-Moberg, K. (1996). Neuroendocrinology of the mother—child interaction. *Trends in Endocrinology & Metabolism*, *7*(4), 126-131. https://doi.org/10.1016/1043-2760(96)00036-7

Van Steensel, R. (2006). Relations between socio-cultural factors, the home literacy environment and children's literacy development in the first years of primary education. *Journal of Research in Reading*, *29*(4), 367-382. doi: 10.1111/j.1467-9817.2006.00301.x

Varner, M. W., & Esplin, M. S. (2005). Current understanding of genetic factors in preterm birth. *BJOG: An International Journal of Obstetrics & Gynaecology*, *112*(s1), 28-31. doi: 10.1111/j.1471-0528.2005.00581.x

Veen, S., Ens-Dokkum, M. H., Schreuder, A. M., Verloove-Vanhorick, S. P., Ruys, J. H., & Brand, R. (1991). Impairments, disabilities, and handicaps of very preterm and very-low-birthweight infants at five years of age. *The Lancet*, *338*(8758), 33-36. https://doi.org/10.1016/0140-6736(91)90015-H

Verduci, E., Banderali, G., Barberi, S., Radaelli, G., Lops, A., Betti, F., ... Giovannini, M. (2014). Epigenetic effects of human breast milk. *Nutrients*, *6*(4), 1711-1724. doi:10.3390/nu6041711

Wang, Z. (2013). A moderated transactional link between child behavioral problems and parenting: A longitudinal-and behavioral-genetic study (Doctoral dissertation). Virginia

Polytechnic Institute and State University, Blacksburg, Virginia. http://hdl.handle.net/10919/50824

Wen, S. W., Fung Kee Fung, K., Huang, L., Demissie, K., Joseph, K. S., Allen, A. C., & Kramer, M. S. (2005). Fetal and neonatal mortality among twin gestations in a Canadian population: The effect of intrapair birthweight discordance. *American journal of perinatology*, *22*(5), 279-286. doi: 10.1055/s-2005-870899

Whitaker, A. H., Feldman, J. F., Van Rossem, R., Schonfeld, I. S., Pinto-Martin, J. A., Torre, C., ... & Paneth, N. S. (1996). Neonatal cranial ultrasound abnormalities in low birth weight infants: relation to cognitive outcomes at six years of age. *Pediatrics*, *98*(4), 719-729.

Whitelaw, A., Heisterkamp, G., Sleath, K., Acolet, D., & Richards, M. (1988). Skin to skin contact for very low birthweight infants and their mothers. *Archives of disease in childhood*, *63*(11), 1377-1381. http://dx.doi.org/10.1136/adc.63.11.1377

Wiig, E. H., & Secord, W. A. (1989). *Test of Language Competence-Expanded Edition*. San Antonio, TX, USA: Pearson.

Wong, A. H., Gottesman, I. I., & Petronis, A. (2005). Phenotypic differences in genetically identical organisms: the epigenetic perspective. *Human molecular genetics*, *14*(suppl 1), R11-R18. https://doi.org/10.1093/hmg/ddi116

Wong, C. C. Y., Meaburn, E. L., Ronald, A., Price, T. S., Jeffries, A. R., Schalkwyk, L. C., ... & Mill, J. (2014). Methylomic analysis of monozygotic twins discordant for autism spectrum disorder and related behavioural traits. *Molecular psychiatry*, *19*(4), 495-503. doi:10.1038/mp.2013.41

World Health Organization. (2001, May 1). *The World Health Organization's infant feeding recommendation*. Retrieved January 23, 2017, from http://www.who.int/nutrition/topics/infantfeeding_recommendation/en/index.html

World Health Organization. (2009). *Infant and young child feeding: model chapter for textbooks for medical students and allied health professionals.* Geneva, Switzerland: WHO Press.

Zimmerman, F. J., Christakis, D. A., & Meltzoff, A. N. (2007). Associations between media viewing and language development in children under age 2 years. *The Journal of pediatrics*, *151*(4), 364-368. http://dx.doi.org/10.1016/j.jpeds.2007.04.071

Zubrick, S. R., Taylor, C. L., Rice, M. L., & Slegers, D. W. (2007). Late language emergence at 24 months: An epidemiological study of prevalence, predictors, and covariates. *Journal of Speech, Language, and Hearing Research*, *50*(6), 1562-1592. doi:10.1044/1092-4388(2007/106)

CHAPTER 5:

Conclusion

The present work focused on understanding what contextual factors influence children's communication development and the processes by which they do so. A more detailed understanding of such developmental processes can help shape health-related social policy and improve the quality of life for individuals with neurodevelopmental communication impairments and their families. As highlighted in Chapter 1, the present dissertation was grounded in the dynamic systems theory of development (Lerner, 2006; Oyama, Griffiths, & Gray, 2001; Smith & Thelen, 2003; Thelen, 2005; Thelen & Smith, 2006) and a distributed model of communication (DeThorne & Miller, 2014; Hengst, 2015), both of which emphasize development as a dynamic and context-dependent system that unfolds over time and is shaped by a myriad of causal influences. In the paragraphs that follow I review key findings from Chapters 2 through 4 and how they have contributed to my understanding of communication development and neurodevelopmental communication impairments, and discuss implications and future directions.

Aside from genetics, research exploring the causal influences on communication development has primarily focused on examining the influence of linguistic factors, and has highlighted the need to develop more effective evidence-based interventions, particularly as it relates to maintenance and generalization effects (e.g., Paul, 2008). Chapter 2 of the present dissertation explicitly assessed the effectiveness of a speech-language intervention in six children 2-4 years of age with various neurodevelopmental conditions. Despite the broad range of treatments available for speech and language development, few target the important milestone of multisyllabic productions and even fewer incorporate both speech and language approaches. This intervention study focused on facilitating multisyllabic productions using an integrated speech-language approach in children who were at the single-word stage of development. We provided multimodal input and provided the opportunity for multimodal output using a myriad of supports and strategies. More specifically, the treatment combined a) motor practice using VocSyl, a novel computerized software to provide real-time visual feedback, and b) child-centered developmental play using a mid-tech speech-generating device (GoTalk 20+) to model and elicit productions within naturalistic interactions. Treatment strategies used ranged from tactile and visual articulation cues to phonological simplification of the multisyllabic targets.

By using the same data set included in DeThorne, Aparicio Betancourt, Karahalios, Halle, and Bogue (2015), Chapter 2 highlighted the contributions of within-participant methodologies to understanding individual trajectories and shaping clinical and educational practice. Overall, this study provided converging within-participant evidence for the use of a multimodal, integrated speech-language intervention to facilitate multisyllabic productions, albeit with substantial individual variability. More specifically, the single-case experimental design revealed more conservative estimates of treatment effect compared with the group results presented in DeThorne and colleagues (2015), in which a functional relationship between the intervention and multisyllabic productions was established for only three of the six participants. In support of maintenance and generalization effects, five of the six children were able to maintain or increase their gains on the multisyllabic treated targets based on a five-week post-treatment maintenance session, and all children showed gains in control (untreated) targets, as well as gains in multisyllabic productions more broadly based on pre- and post-intervention communicative development expressive vocabulary inventories. In addition to assessing the effectiveness of the intervention, a second aim of the present study was to address the complementary role of individual science in uncovering knowledge. Single-case designs, a type of within-participant methodology, allow you to assess causal relationships, similar to group designs, while providing a more in-depth characterization of individual children over time, and offering the flexibility to tailor interventions to the needs of individual children. They offer specific advantages for healthcare and educational providers such as increasing the feasibility for the designs to be implemented in applied contexts, and have the potential to better match individual children's profiles with specific interventions.

As a complement to the behavioral intervention study presented in Chapter 2, Chapter 3, recently published in ACM Transactions on Accessible Computing, explored the novel use of noninvasive biosensors to measure electrical conductance across the skin during intervention as a potential support for communication (Aparicio Betancourt, DeThorne, Karahalios, & Kim, 2017). Given the moderate to high comorbidity between neurodevelopmental communication impairments and socioemotional and behavioral difficulties (Qi & Kaiser, 2004; Redmond & Rice, 1998; Toppelberg, Medrano, Morgens, & Nieto-Castañon, 2002), there is interest in exploring the role of technology in supporting communication practices involving children with neurodevelopmental disabilities. Skin conductance has been widely used as a marker for emotional arousal; it is mediated by sympathetic cholinergic sudomotor nerve fibers stemming mainly from

the ipsilateral limbic hypothalamic source (Benedek and Kaernbach 2010; Boucsein, 2012). However, traditionally its use has been limited to highly controlled settings, with restricted external validity. Nonetheless, recent technological advancements have led to the introduction of ambulatory biosensors, making *in situ* studies of electrodermal activity (EDA) more feasible. Chapter 3 specifically draws on user experience from two interventions, the speech-language therapy described in Chapter 2 and occupational therapy (Snodgrass et al., 2015), to explore the benefits of *in situ* skin conductance assessment in eight children, ages 2-11, with neurodevelopmental communication impairments. During the speech-language intervention (Aparicio Betancourt, DeThorne, & Karahalios, 2013), we examined the association between *in situ* skin conductance and off-line behavioral coding of emotional valence (i.e., affect), as well as the association among different skin conductance measures, as area under the curve is relatively unexplored. Skin conductance response frequency was higher for high positive and high negative emotional valence ratings compared to neutral ratings, providing evidence of concurrent validity for *in situ* skin conductance assessment, albeit with a small effect size (i.e., 2.3%). In regard to associations between skin conductance measures, significant moderate to high correlations were observed. Specific to skin conductance level area under the curve, findings suggested this measure may be mediated by similar psychophysiological or neurological sources than skin conductance response frequency.

During the occupational therapy intervention reported in Chapter 3, we examined the effects of a pressure vest on academic engagement, challenging behaviors, and skin conductance as children with neurodevelopmental disabilities often have difficulties with sensory integration, and pressure vests are commonly used within occupational therapy as a means to help regulate physiological arousal. Use of the pressurized vest however, was not associated with skin conductance or either of the two behavioral measures. Preliminary guidelines focused on behavioral and technical findings associated with the use of biosensors to measure skin conductance *in situ* across both the speech-language and occupational therapy studies. To illustrate, we recommended specific strategies to assist with desensitization to novel stimuli, such as the use of a social story, and provided recommended thresholds to use for data analysis. Our aim was to contribute to the development of reliable noninvasive technology for measuring emotional arousal via skin conductance in applied settings, by highlighting that although it is feasible to record EDA *in situ,* both behavioral and technical challenges remain. We concluded

that skin conductance assessment holds qualified promise in facilitating communication with the hope of building better environments to support the learning and social interaction of children with neurodevelopmental communication impairments.

Finally, Chapter 4 specifically addressed the need to explore contextual influences on communication development, beyond linguistic factors. More specifically, the work in this chapter emerged from our prior work which indicated a need to study a) a broader range of environmental influences on child language development while acknowledging the complex interplay between genetic and environmental influences (Rogers, Nulty, Aparicio Betancourt, & DeThorne, 2015); and b) the source of nonshared environmental (NSE) influences on language that increased with age and potentially differed based on assessment context (Harlaar, DeThorne, Mahurin-Smith, Aparicio Betancourt, & Petrill, 2016). This chapter used a longitudinal MZ twin difference approach, a genetically-sensitive design, to examine candidate nonshared environmental influences on language development. We assessed the extent to which discordance in specific environmental variables (i.e., birthweight, home reading exposure at age 7, and parenting at age 7) was associated with discordance in Productive and Formal language at mean ages 10 and 12 years. Although we originally intended to examine breastfeeding as an environmental variable as well, it did not demonstrate the requisite MZ discordance needed for such analyses. For the remaining environmental variables considered, significant effects emerged for birthweight, home reading exposure, and self-reported parenting, primarily in association with Formal language at the later time point, with effect sizes ranging from 4% to 20%. NSE effects were also moderated by the extent of MZ twin language discordance, with effects increasing with increased language discordance. To the extent that the effects are causal, this work has implications for the future of prenatal care, access to early intervention for children born low birthweight, and development of subsidized culturally-atuned programs that aim to reduce parent stress, increase family access to literacy-based materials/practices, and strengthen parent-child relationships. Findings also underscore the need to incorporate multiple forms of assessment in clinical and educational settings.

Altogether, my dissertation work highlights the need to study a wide array of contextual factors influencing communication development, the need to incorporate diverse and complementary methodologies, and the need for the development of effective communication supports for children with neurodevelopmental communication impairments. Communication permeates every aspect

of life, from academic achievement to mental health, and it is particularly relevant to understand communication development within a multicausal, dynamic, and context-dependent framework that acknowledges the complex interplay between genetic and environmental influences. Future work within the intersection of communication science and neuroscience should prioritize uncovering candidate environmental influences on communication development and their associated mechanisms in order to develop more effective and individualized interventions and supports.

References

Aparicio Betancourt, M., DeThorne, L. & Karahalios, K. (2013, November). *Facilitating multisyllabic productions & assessing sympathetic arousal in children with developmental disorders*. Poster presented at the Society for Neuroscience Conference, San Diego, California.

Aparicio Betancourt, M., DeThorne, L. S., Karahalios, K. & Kim, J. G. (2017). Skin conductance as an *in situ* marker for emotional arousal in children with neurodevelopmental communication impairments: Methodological considerations and clinical implications. *ACM Transactions on Accessible Computing*, *9*(3), 8. http://dx.doi.org/10.1145/3035536

Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*, *190*(1), 80-91. doi: 10.1016/j.jneumeth.2010.04.028.

Boucsein, W. (2012). *Electrodermal activity* (2nd ed.). New York, NY, USA: Springer Science & Business Media.

DeThorne, L., Aparicio Betancourt, M., Karahalios, K., Halle, J., & Bogue, E. (2015). Visualizing syllables: Real-time computerized feedback within a speech–language intervention. *Journal of Autism and Developmental Disorders*, *45*(11), 3756-3763. doi:10.1007/s10803-014-2274-8

DeThorne, L.S., & Miller, C. A. (2014). Communication development: Distributed across people, resources and time. In J. Nussbaum (Ed.), *Handbook of lifespan communication* (53-70). New York, New York, USA: Peter Lang Publishing

Harlaar, N., DeThorne, L. S., Mahurin-Smith, J., Aparicio Betancourt, M., & Petrill, S. A. (2016). Longitudinal Effects on Early Adolescent Language: A Twin Study. *Journal of Speech, Language, and Hearing Research*, *59*(5), 1059-1073. doi:10.1044/2016_JSLHR-L-15-0257

Hengst, J. A. (2015). Distributed communication: Implications of cultural-historical activity theory (CHAT) for communication disorders. *Journal of communication disorders*, *57*, 16-28. http://dx.doi.org/10.1016/j.jcomdis.2015.09.001

Lerner, R. M. (2006). Developmental science, developmental systems, and contemporary theories of human development. In W. Damon & R.M. Lerner (Eds.), *Handbook of Child Psychology* (Vol.1), (1-17). Hoboken, NJ, USA: John Wiley & Sons. doi:10.1002/9780470147658.chpsy0101

Oyama, S., Griffiths, P. E., & Gray, R. D. (2001). Introduction: What is developmental systems theory. In S. Oyama, P. E. Griffiths, & R. D. Gray (Eds.), *Cycles of contingency: Developmental systems and evolution* (pp. 1-11). Cambridge, Massachusetts, USA: MIT Press.

Paul, R. (2008). Interventions to improve communication. *Child and Adolescent Psychiatric Clinics of North America*, *17*(4), 835–x. http://doi.org/10.1016/j.chc.2008.06.011

Qi, C. H., & Kaiser, A. P. (2004). Problem behaviors of low-income children with language delays: An observation study. *Journal of Speech, Language, and Hearing Research,* 47(3), 595-609. doi: 10.1044/1092-4388(2004/046).

Redmond, S. M., & Rice, M. L. (1998). The socioemotional behaviors of children with SLI: Social adaptation or social deviance? *Journal of Speech, Language, and Hearing Research,* 41(3), 688-700. doi: 10.1044/jslhr.4103.688.

Rogers, C. R., Nulty, K. L., Aparicio Betancourt, M., & DeThorne, L. S. (2015). Causal effects on child language development: A review of studies in communication sciences and disorders. *Journal of communication disorders*,*57*, 3-15. http://dx.doi.org/10.1016/j.jcomdis.2015.06.004

Smith, L. B., & Thelen, E. (2003). Development as a dynamic system. *Trends in cognitive sciences*, *7*(8), 343-348. http://dx.doi.org/10.1016/S1364-6613(03)00156-6

Snodgrass, M. R., Kim, J. G., Karahalios, K., Borden, M., & Halle, J. W. (2015). *The effects of a pressure vest on task engagement, challenging behavior, and a physiological measure of stress for a child with intellectual disability*. University of Illinois Digital Environment for Access to Learning and Scholarship. http://hdl.handle.net/2142/79051.

Thelen, E. (2005). Dynamic systems theory and the complexity of change. *Psychoanalytic Dialogues*, *15*(2), 255-283. http://dx.doi.org/10.1080/10481881509348831

Thelen, E., & Smith, L. B. (2006). Dynamic systems theories. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology* (Vol. 1), (258-312). Hoboken, NJ, USA: John Wiley & Sons.

Toppelberg, C. O., Medrano, L., Morgens, L. P., & Nieto-Castanon, A. (2002). Bilingual children referred for psychiatric services: associations of language disorders, language skills, and psychopathology. *Journal of the American Academy of Child & Adolescent Psychiatry*, *41*(6), 712-722. https://doi.org/10.1097/00004583-200206000-00011

APPENDIX A:

Social Story Akin to That Used as a Desensitization Strategy during the SL-EDA Study

I get to come to the Speech Building to play.



I can wait in the lobby.



Mariana will come get me.



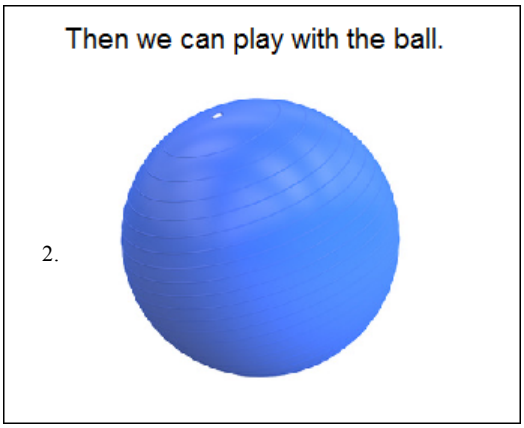
First, we will put the bracelets on my ankles.



1.

1. Affectiva, 2012

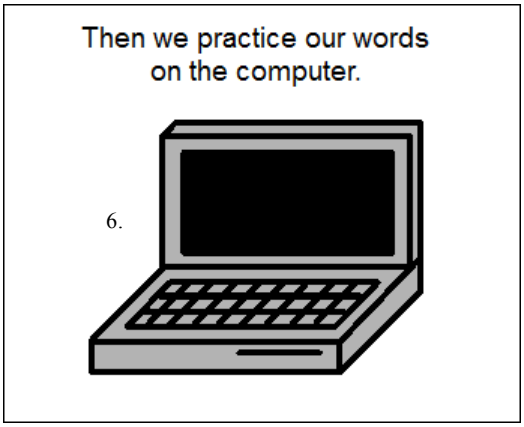
The Picture Communication Symbols ©1981–2015 by Mayer-Johnson LLC a Tobii Dynavox company. All Rights Reserved Worldwide. Used with permission. Boardmaker® is a trademark of Mayer-Johnson LLC.



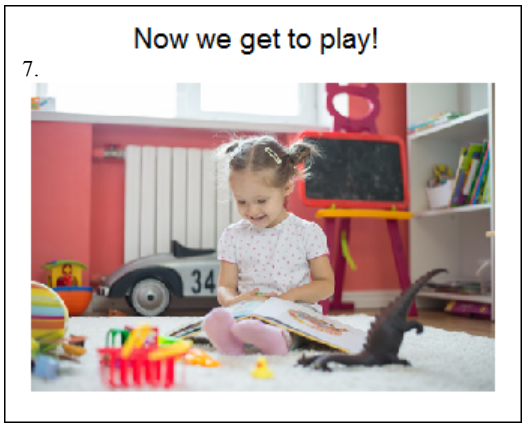
2. © Alexlmx / Fotolia



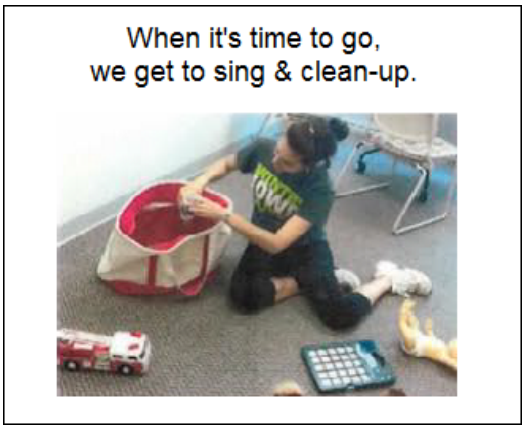
3. © Ulzanna / Fotolia
4. © Mayer-Johnson / Boardmaker®
5. © Mayer-Johnson / Boardmaker®

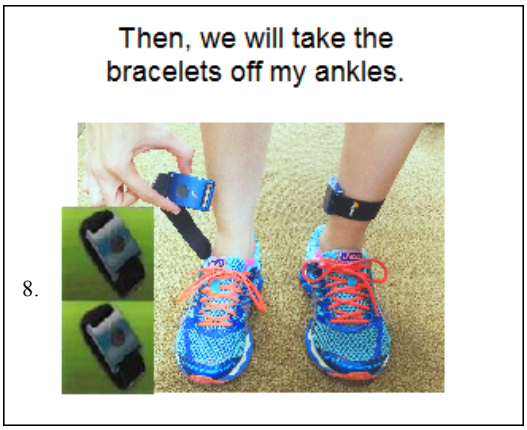


6. © Mayer-Johnson / Boardmaker®



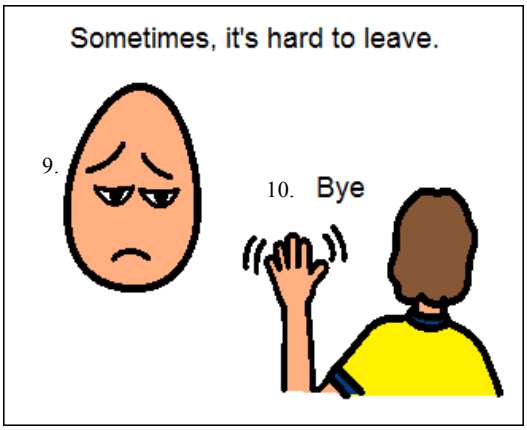
7. © Rasstock / Fotolia



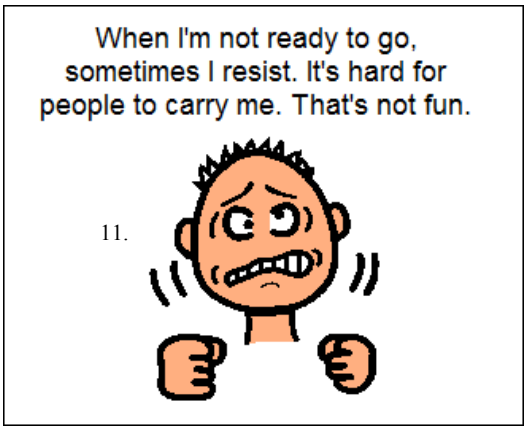


8. Affectiva, 2012

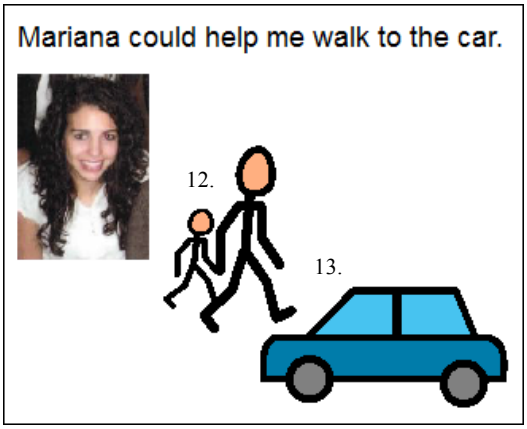
The Picture Communication Symbols ©1981–2015 by Mayer-Johnson LLC a Tobii Dynavox company. All Rights Reserved Worldwide. Used with permission. Boardmaker® is a trademark of Mayer-Johnson LLC.



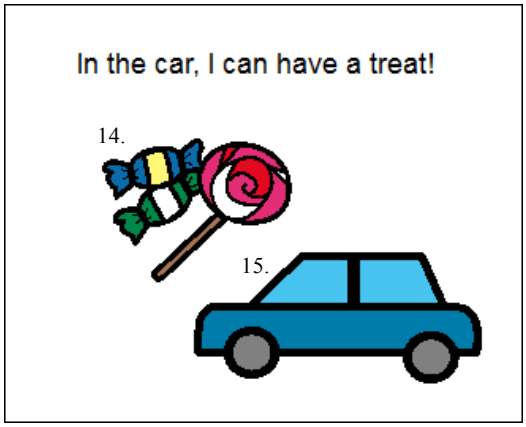
9. © Mayer-Johnson / Boardmaker®
10. © Mayer-Johnson / Boardmaker®



11. © Mayer-Johnson / Boardmaker®

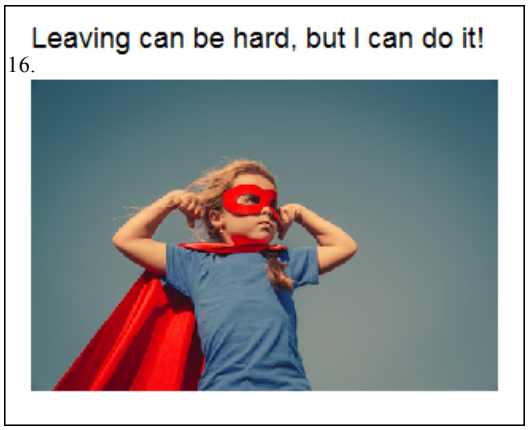


12. © Mayer-Johnson / Boardmaker®
13. © Mayer-Johnson / Boardmaker®



14. © Mayer-Johnson / Boardmaker®
15. © Mayer-Johnson / Boardmaker®





16. © Yuryimaging / Fotolia

The Picture Communication Symbols ©1981–2015 by Mayer-Johnson LLC a Tobii Dynavox company. All Rights Reserved Worldwide. Used with permission. Boardmaker® is a trademark of Mayer-Johnson LLC.

APPENDIX B:

PARCHISY Cooperative-Play Tasks

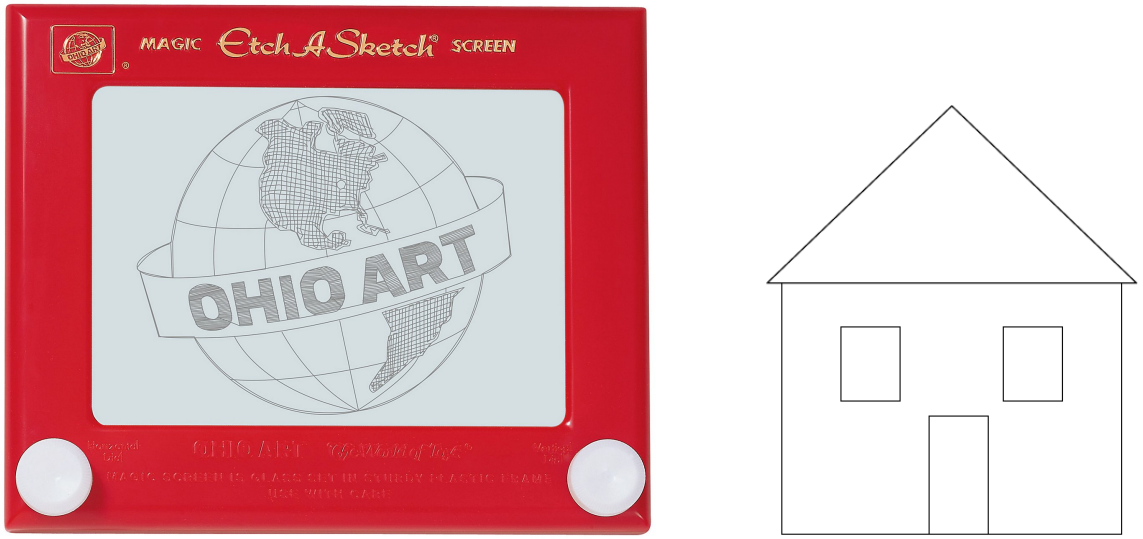


Figure B.1: Ohio Art Etch A Sketch & Etch A Sketch Stimulus



Figure B.2: Example of a marble maze game

APPENDIX C:

MZ Difference Correlations in Language Outcomes and Environmental Measures for the Full Samples and for Subsamples of Boys and Girls.

Of the 72 total correlation coefficients, fifteen reached significance and one reached marginal significance, which is greater than the four significant correlations that would be expected by chance.

Table C.1. *Correlations between MZ differences in language at mean age 10 (HV5) and MZ differences in environment.*

Environmental Measures	Formal Language	Productive Language
Birthweight All	- 0.02	0.02
Boys	0.16	- 0.06
Girls	- 0.17	0.11
Parenting at 7		
PFQ-Positivity All	0.13	0.02
Boys	**0.34***	- 0.26
Girls	- 0.07	0.20
PFQ-Negativity All	**- 0.45****	- 0.09
Boys	**- 0.34***	0.23
Girls	**- 0.52****	**- 0.30***
PARCHISY-Positivity All	0.03	0.01
Boys	0.01	- 0.19
Girls	- 0.01	0.09
PARCHISY-Negativity All	0.05	- 0.02
Boys	0.10	0.13
Girls	0.05	- 0.09
Home Reading Exposure at 7 All	0.12	- 0.15
Boys	**0.37***	- 0.21
Girls	0.01	- 0.07

Note. One-tailed *p*-values: *$p<.05$; **$p<.01$. All: N = 61-113 pairs. Boys/Girls: N = 25-67 pairs. Boy/girl distinction is based on binary parent identification of "sex" with options "girl" or "boy", thereby conflating sex and gender. PFQ = Parent Feelings Questionnaire; PARCHISY = Parent-Child Interaction System.

Table C.2. *Correlations between MZ differences in language at mean age 12 (HV7) and MZ differences in environment.*

Environmental Measures	Formal Language	Productive Language
Birthweight All	**0.23****	**0.22***
Boys	0.19	**0.39***
Girls	**0.29***	0.16
Parenting at 7		
PFQ-Positivity All	**0.19+**	- 0.08
Boys	0.29	- 0.11
Girls	0.13	- 0.09
PFQ-Negativity All	**- 0.45****	0.01
Boys	**- 0.47****	- 0.08
Girls	**- 0.45****	0.08
PARCHISY-Positivity All	- 0.07	- 0.04
Boys	- 0.06	- 0.29
Girls	- 0.06	0.04
PARCHISY-Negativity All	0.18	0.04
Boys	0.29	0.30
Girls	0.10	- 0.06
Home Reading Exposure at 7 All	**0.21***	0.01
Boys	**0.53***	0.28
Girls	0.07	- 0.05

Note. One-tailed p-values: +p=.05; *p<.05; **p<.01. All: N = 57-105 pairs. Boys/Girls: N = 21-65 pairs. Boy/girl distinction is based on binary parent identification of "sex" with options "girl" or "boy", thereby conflating sex and gender. PFQ = Parent Feelings Questionnaire; PARCHISY = Parent-Child Interaction System.