Episodic Detail Production and Semantic Coherence in Down Syndrome and Fragile X Syndrome: Longitudinal Findings from Expressive Language Sampling

Hailey Van Vorce

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

Master of Science
In
Psychology

Jamie O. Edgin
Rachel A. Diana
Rosanna P. Breaux

December 9, 2025
Blacksburg, Virginia

Keywords: Autobiographical memory, down syndrome, fragile X syndrome, Expressive Language Sampling, narrative coherence

Episodic Detail Production and Semantic Coherence in Down Syndrome and Fragile X Syndrome: Longitudinal Findings From Expressive Language Sampling

Hailey Van Vorce

ABSTRACT

Autobiographical memory requires the integration of episodic and semantic information and is closely tied to expressive language abilities. This study examined episodic detail production and narrative coherence in children and adolescents with Down syndrome (DS) and Fragile X syndrome (FXS) using conversational samples from the Expressive Language Sampling (ELS) Conversation task (Abbeduto et al., 2020, 2023). Participants ($N = 50$) contributed one matched autobiographical topic at two visits approximately 18 months apart. Episodic and semantic details were coded using the Autobiographical Interview (AI) framework (Levine et al., 2002), and narrative coherence was assessed using Semantic Distance (SemDis), a computational measure of conceptual relatedness (Beaty & Johnson, 2021). Multilevel models evaluated whether diagnostic group, expressive language, narrative length, and time predicted autobiographical memory performance. Across aims, children showed substantial variability in narrative output, with greater within-group than between-group differences. Diagnostic group did not significantly predict episodic detail production, and episodic content showed minimal change across time. Word count was the only significant predictor, indicating that children who produced more language provided more episodic content. No demographic or language variables uniquely predicted episodic detail production once narrative length was controlled. Semantic coherence was also stable across visits and did not differ by diagnostic group or narrative length. The only significant effect was a diagnostic group × CELF-FS interaction: higher expressive syntax predicted more coherent narratives among children with DS, whereas children with FXS showed a slight decrease in coherence as expressive syntax increased. Overall, findings indicate that expressive output, rather than diagnostic status, is the primary driver of autobiographical narrative performance in DS and FXS.

Episodic Detail Production and Semantic Coherence in Down Syndrome and Fragile X Syndrome: Longitudinal Findings From Expressive Language Sampling

Hailey Van Vorce

GENERAL AUDIENCE ABSTRACT

Autobiographical memory is the ability to talk about personal past experiences. It supports communication, relationships, and the sharing of meaningful events in everyday life. This study explored how children and adolescents with Down syndrome and Fragile X syndrome describe personal experiences during natural conversation. We focused on two parts of their storytelling: how many specific details they shared about an event and how clearly their ideas stayed connected to the topic. Fifty participants completed the same conversation task at two points in time that were about eighteen months apart. We counted the details they provided and used a computer-based method to measure how organized and on topic their narratives were. Children in both groups showed a wide range of storytelling abilities. The two groups were similar overall, and their performance did not change very much across the two visits. The strongest predictor of how many details a child shared was simply how much they talked. Narrative organization was mostly stable and did not differ between groups, although expressive language ability played a somewhat different role in each group. Overall, these findings suggest that individual differences in how much children speak have a greater influence on their autobiographical storytelling than their specific diagnosis.

Acknowledgements

I would like to express my sincere appreciation to the individuals and teams who made this project possible. I am deeply grateful to research assistants, Madison Fleming and Alyssa An, for their careful work, attention to detail, and consistent support throughout the coding and data preparation phases. I also extend my heartfelt thanks to Chelsea Armour, whose contributions were invaluable. Chelsea provided extensive support with coding, troubleshooting, conceptual decisions, and overall guidance across multiple stages of the project, and her generosity and expertise strengthened this work in countless ways.

I am especially thankful to my advisor, Jamie Edgin, for her mentorship, patience, and unwavering encouragement. Her insight and support have shaped every part of this project and my development as a researcher.

I also extend my appreciation to the team at the University of California Davis, including Angela Thurman and Leonard Abbeduto, for granting access to the Expressive Language Sampling (ELS) dataset and transcripts that formed the foundation of this study. This study represents a secondary analysis of transcripts and data collected as part of the ELS project. The development and testing of ELS were supposed by grants from the National Institutes of Health (R01HD074346, U54HD079125, UL1TR001860, P50HD103526, R01DC012033) and the Simons Foundation (SFARI 383668).

Finally, I would like to thank the Levine Lab for providing the training and practice materials used for Autobiographical Interview coding. Their resources and scoring framework were essential to the rigor and reliability of this work. I am truly grateful to all who contributed their time, expertise, and collaboration to this project.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

Chapter 1. INTRODUCTION

Memory is a central cognitive process that supports personal identity, communication, and social connections. Two major systems, episodic and semantic memory, shape how individuals recall and interpret their experiences. Episodic memory refers to the recollection of personally experienced events that occur at a specific time and place and can include sensory, emotional, and contextual information (Tulving, 1972). These memories allow individuals to describe moments from their lives and to articulate what happened, how it felt, and why it mattered. Semantic memory, in contrast, involves general knowledge about the world, including facts, concepts, and meanings that can be accessed without reference to the context in which they were learned (Tulving, 1985). Although the two systems are distinct, they often support and enrich one another. Episodic experiences contribute to the accumulation of semantic knowledge, and semantic frameworks provide conceptual scaffolding that helps organize and retrieve episodic details (Grilli & Verfaellie, 2014; Grilli et al., 2018).

This interplay is foundational for autobiographical memory, which requires individuals to blend event-specific details with broader personal and world knowledge (Barton-Husley et al., 2017). Autobiographical memory is essential for social relationships, emotional well-being, academic engagement, and the communication of personal experiences across both familiar and unfamiliar contexts (Boudreau & Chapman, 2000; Chapman & Hesketh, 2000; Wang et al., 2020). Critically, autobiographical memory underlies the ability to recount past experiences in any setting, from everyday conversations to clinical interactions and other evaluative contexts. For individuals with neurodevelopmental conditions such as fragile X syndrome (FXS) and Down syndrome (DS), these abilities often develop differently. Research consistently shows that individuals in these groups tend to produce narratives with fewer elaborated episodic details and internal-state references, while relying more on general or repeated information (Abbeduto et al., 2007; Finestack et al., 2012; McLennan et al., 2011). These narrative tendencies influence daily communication, learning, therapeutic engagement, and how individuals are perceived in contexts requiring detailed recall (Cornish et al., 2007; Vicari, 2001).

Although the current study does not directly examine forensic interviewing, this broader context is important to acknowledge. Individuals with intellectual and developmental disabilities, including those with FXS and DS, are disproportionately likely to experience victimization, yet differences in communication and memory profiles often lead others to underestimate the credibility of their accounts (Harrell, 2021; Peled et al., 2004). Because autobiographical memory is central to recounting lived experiences, understanding how episodic and semantic memory develop in these groups can inform the kinds of supports, expectations, and question types that facilitate more accurate and comfortable recall in clinical, educational, and investigative settings. This background also provides a rationale for later methodological decisions, including the types of questions analyzed within the Autobiographical Interview (AI) coding scheme.

FXS is the most common inherited cause of intellectual disability, occurring in approximately 1 in 4,000 males and 1 in 8,000 females (Hunter et al., 2014). It is caused by a mutation in the FMR1 gene on the X chromosome that reduces or eliminates the production of FMRP, a protein essential for synaptic plasticity and neural development (Kremer et al., 1993). Individuals with FXS frequently demonstrate challenges in working memory, executive functioning, expressive language, and narrative organization, all of which influence the ability to generate and structure episodic information (Gross et al., 2011; Barton-Husley et al., 2017). Despite these challenges, semantic memory is often relatively preserved (McLennan et al., 2011). Behavioral characteristics associated with FXS, including social anxiety and attentional dysregulation, may further shape autobiographical recall and the amount of detail provided (Abbeduto et al., 2020; Crane et al., 2020).

Like individuals with FXS, those with Down syndrome also show distinct language and memory profiles that influence autobiographical expression. DS occurs in approximately 1 in 700 live births (Mai et al., 2019) and results from an additional full or partial copy of chromosome 21 (Sherman et al., 2007). Individuals with DS often demonstrate strengths in receptive vocabulary and general semantic knowledge, paired with challenges in expressive language, verbal short-term memory, and syntactic production (Chapman & Hesketh, 2000; Jarrold et al., 2002). Narrative research indicates that, although individuals with DS often produce stories with recognizable global structure, their narratives frequently contain fewer elaborated episodic details, fewer internal-state references, and less integration across story elements (Boudreau & Chapman, 2000; Miles & Chapman, 2002; Segal & Pesco, 2015). As a result, narratives may rely more heavily on generalized semantic information and contain fewer specific contextual features.

Given these developmental profiles, it is important to examine both strengths and variability in episodic and semantic memory in individuals with FXS and DS. Prior work on memory in these groups has relied heavily on structured memory tasks, which, while informative, do not fully capture how individuals communicate about their lives in natural contexts (Abbeduto et al., 2023; Finestack et al., 2012). Understanding autobiographical memory requires examining how individuals spontaneously describe personally meaningful experiences.

ELS provides an ideal way to address this limitation. ELS elicits spontaneous speech through semi-structured conversations about familiar topics (Abbeduto et al., 2023). The procedures were originally developed by Abbeduto and colleagues (2023) in a large multi-site research study focused on language and communication in neurodevelopmental disorders. As a result, the transcripts available for this thesis represent an exceptionally rich, ecologically valid, and psychometrically robust dataset. Although the ELS study was not specifically designed to examine memory, its open-ended conversational prompts naturally elicit both episodic and semantic content, making the transcripts well suited for analyzing autobiographical memory. The strength and rigor of the original study by Abbeduto et al. (2023) allows the present project to conduct a meaningful secondary analysis, building on a dataset that was collected with high fidelity and standardized procedures across sites.

To quantify memory content within these narratives, the present study applies the AI coding scheme (Levine et al., 2002). The AI distinguishes episodic (internal) details, which describe specific events, from semantic (external) details, including general knowledge, personal facts, and repeated information. While AI coding captures detail-level content, it does not fully characterize how ideas are conceptually related within the narrative. To address this, the study integrates semantic distance analysis using SemDis (Beaty & Johnson, 2021), a computational method that measures conceptual relatedness across narrative elements. Episodic narratives often display tightly connected conceptual clusters, whereas narratives that rely more heavily on semantic information may show broader conceptual spread. Combining AI coding with semantic distance analysis therefore offers a more comprehensive view of both the content and conceptual organization of autobiographical memory.

With extensive validation across multiple phases of research, ELS has proven to be a reliable and practical approach for assessing language skills in individuals with intellectual and developmental disabilities. This places the current study in a uniquely strong position, as it leverages high-quality transcripts collected through a rigorous standardized protocol by Abbeduto et al. (2023). Although this project represents a secondary analysis, the richness and methodological strength of the original ELS study provide an ideal foundation for examining the development and differentiation of autobiographical memory in FXS and DS.

The present study examines the development and differentiation of episodic and semantic autobiographical memory in individuals with FXS and DS across an 18-month period. Two outcome domains are investigated: episodic detail production (Levine AI and narrative coherence, measured through semantic distance.

Aim 1 is to evaluate stability and change in episodic (internal) detail production across two time points and to determine whether individual differences in semantic detail production, chronological age, cognitive ability, expressive language ability, and diagnostic group predict variability in these trajectories. Based on prior findings, it is expected that individuals with both FXS and DS will produce more semantic than episodic details at both time points. Individuals with stronger cognitive and language abilities, as well as those who are older, are anticipated to demonstrate greater stability or increases in episodic detail production over time.

Aim 2 is to examine developmental change in narrative coherence, indexed through semantic distance, across the same period. This aim also evaluates whether cognitive ability, expressive language skills, age, and diagnostic group are associated with initial coherence levels and their change over time. Coherence is expected to improve for both groups, with individuals who have stronger language and cognitive abilities showing the greatest gains. Individuals with FXS may demonstrate more consistent semantic coherence across time, whereas individuals with DS may show greater variability due to broader challenges integrating memory and language during narrative production.

Chapter 2. METHODS

2.1 Design

This study uses a longitudinal secondary data design to examine autobiographical memory in individuals with DS and FXS. Narrative samples were drawn from the ELS Conversation task, originally collected as part of a multi-site study conducted by Abbeduto and colleagues (Abbeduto et al., 2020; Abbeduto et al., 2023). For the present project, two time points spaced approximately 18 months apart were selected to allow for the examination of developmental changes in narrative memory.

The primary focus of this study is the differentiation and development of episodic and semantic autobiographical memory over time. The primary outcome variables were: (a) the total number of episodic (internal) details and semantic (external) details produced within each narrative, and (b) the semantic similarity between each participant's narrative and the corresponding examiner prompt. Episodic and semantic detail counts were derived using the standard AI scoring protocol (Levine et al., 2002), applied to highlighted episodic transcript segments. Narrative coherence was assessed using semantic distance metrics generated by the archived SemDis system, which quantifies the conceptual alignment between a participant's narrative responses and the conversational prompts provided by the examiner.

Multilevel modeling (MLM) was employed to account for the nested structure of the data, with repeated narrative observations (Time 1 and Time 3) nested within individuals. This framework allowed for the examination of both within-person change across time points and between-group differences (DS vs. FXS) in autobiographical memory trajectories. Between-person predictors included diagnostic group, sex, chronological age, and expressive language ability (as indexed by the Clinical Evaluation of Language Fundamental [CELF-4] Formulated

Sentences [FS] raw score). Child word count from the highlighted segments of the transcripts were included as a covariate in both multilevel models to adjust for individual differences in verbal output. This analytic structure supported the investigation of how cognitive and linguistic factors contribute to changes in episodic detail production and semantic coherence over time.

2.2 Participants

Participants were drawn from a larger longitudinal ELS study conducted by Abbeduto and colleagues (Abbeduto et al., 2020; Abbeduto et al., 2023). Only participants with genetically confirmed DS or FXS and those who completed both visits were included in the present project. Participants ranged in age from 6 to 23 years, with a mean age of 14.39 years (SD = 4.02) at Time 1. All individuals demonstrated intellectual functioning within the intellectual disability range (full scale intelligence quotient [FSIQ] < 70). The sample was 70% male overall, with sex distributions differing between diagnostic groups (91.3% male in FXS and 51.9% male in DS). This difference was statistically significant, $\chi^2(1) = 9.21$, $p = .002$. This pattern is consistent with known male-skewed prevalence of FXS and was therefore included as a covariate in all multilevel models.

Table 1
Participant Characteristics by Diagnostic Group

Variable	FXS (n = 23)	DS (n = 27)	Total (N = 50)	Test Statistic	*p*
Sex, *n* (% male)	21 (91.3%)	14 (51.9%)	35 (70.0%)	$\chi^2(1) = 9.21$	0.002
Chronological Age (years)	M = 14.52 (SD = 4.18)	M = 14.27 (SD = 3.92)	M = 14.39 (SD = 4.02)	$t(48) = 0.23$	0.82
SB-5 Full Scale IQ (FSIQ)	M = 46.14 (SD = 7.09)	M = 45.76 (SD = 6.15)	M = 45.94 (SD = 6.57)	$t(45) = 0.20$	0.85
SB-5 Nonverbal IQ (NVIQ)	M = 47.83 (SD = 7.44)	M = 46.41 (SD = 6.82)	M = 47.07 (SD = 7.08)	$t(45) = 0.65$	0.52

SB-5 Verbal IQ (VIQ)	M = 44.22 (SD = 8.02)	M = 44.89 (SD = 6.91)	M = 44.58 (SD = 7.39)	$t(45) = 0.29$	0.77

Note. Sex differed significantly between diagnostic groups, with a higher proportion of males in the FXS group (91.3%) than the DS group (51.9%) a chi-square test of independence indicated $\chi^2(1) = 9.21$, $p = 0.002$.

Table 2
Household Income Grouped Below vs. Above $75,000 by Diagnostic Group

Income Group	FXS *n* (%)	DS *n* (%)	Total *n* (%)
< $75,000	9 (39.1%)	10 (38.5%)	19 (38.8%)
≥ $75,000	14 (60.9%)	16 (61.5%)	30 (61.2%)
	23	26	49

Note. Income was coded using available data from either time point (Time 1 or Time 3). Codes 1–3 (<$75,000) and 4–7 (≥$75,000) were used for classification. One participant was missing income at both time points. A chi-square test indicated no significant difference in income distribution between diagnostic groups, $\chi^2(1) = 0.00$, p = 1.00.

Racial, ethnic, and socioeconomic characteristics were broadly similar across diagnostic groups. The sample was predominantly White/European American (66%), followed by Multiracial (12%), Hispanic/Latine (10%), African American (6%), Asian/Pacific Islander (2%), and another racial-ethnic identity than those listed (4%). Differences in racial-ethnic distribution approached but did not reach statistical significance (see Table 4). Maternal education ranged from K–8 schooling to a master's degree or higher, and household income spanned the full range of categories assessed (less than $25,000 to more than $250,000); neither variable significantly differed between diagnostic groups (see Tables 2–3).

Table 3
Maternal Caregiver Education: Bachelor's Degree or Higher by Diagnostic Group

Education Level	FXS *n* (%)	DS *n* (%)	Total *n* (%)
Less than bachelor's	11 (47.8%)	13 (48.1%)	24 (48.0%)
Bachelor's degree or higher	12 (52.2%)	14 (51.9%)	26 (52.0%)
	23	27	50

Note. Education categories were collapsed into two groups: less than a bachelor's degree (K–8, some high school, high school graduate/GED, some college, associate/technical degree) and bachelor's degree or higher (BA/BS, some graduate work, master's or higher).
A chi-square test comparing the distribution across groups was not statistically significant, $\chi^2(1) = 0.00$, $p = .985$.

Table 4
Race / Ethnicity by Diagnostic Group (Collapsed Categories)

Race / Ethnicity	FXS *n* (%)	DS *n* (%)	Total *n* (%)
White (non-Hispanic)	15 (65.2%)	18 (66.7%)	33 (66.0%)
Non-White racial/ethnic groups	8 (34.8%)	9 (33.3%)	17 (34.0%)
	23 (100%)	27 (100%)	50 (100%)

Note. Race/ethnicity categories were collapsed into two groups (White/non-Hispanic and non-White racial/ethnic groups) due to small cell sizes in individual race categories. The non-White racial/ethnic group included participants identifying as African American (FXS n = 3, DS n = 0), Asian/Pacific Islander (FXS n = 1, DS n = 0), Hispanic/Latino/a (FXS n = 1, DS n = 4), Other (FXS n = 2, DS n = 0), and Multiracial (FXS n = 1, DS n = 5), for a total of 17 participants. A chi-square test indicated no significant association between race/ethnicity and diagnosis, $\chi^2(1, N = 50) = 0.01$, p = .91.

To evaluate whether the analytic subsample differed from the broader ELS cohort, demographic and cognitive characteristics were compared between groups (see Table 5). The analytic sample did not meaningfully differ from the full dataset on age, sex distribution, nonverbal cognitive ability, or adaptive behavior, supporting the representativeness of the analytic sample.

Table 5

Comparison of Full ELS Sample and Analytic Subsample on Demographic and Cognitive Variables

Variable	Full ELS Sample ($N = 218$)	Analytic Subsample ($n = 50$)	Test Statistic	p
Diagnosis (FXS / DS)	50.9 % / 49.1 %	46.0 % / 54.0 %	$\chi^2(1, N = 218) = 0.63$	0.428
Sex (Male / Female)	66.7 % / 33.3 %	70.0 % / 30.0 %	$\chi^2(1, N = 218) = 0.18$	0.669
Age – Time 1 (years)	$M = 14.99$ ($SD = 4.94$)	$M = 14.81$ ($SD = 4.80$)	$t(216) = 0.23$	0.818
Age – Time 3 (years)	$M = 16.80$ ($SD = 4.96$)	$M = 16.88$ ($SD = 4.69$)	$t(193) = –0.10$	0.920
Full-Scale IQ (FSIQ)	$M = 45.60$ ($SD = 9.26$)	$M = 45.94$ ($SD = 6.54$)	$t(206) = –0.24$	0.814
Nonverbal IQ (NVIQ)	$M = 47.90$ ($SD = 8.72$)	$M = 49.17$ ($SD = 7.52$)	$t(206) = –0.90$	0.367
Verbal IQ (VIQ)	$M = 48.09$ ($SD = 9.70$)	$M = 47.71$ ($SD = 5.98$)	$t(210) = 0.26$	0.799

Note. No significant differences were found between the full ELS dataset and the analytic subsample across demographic or cognitive measures, suggesting that the analytic sample is representative of the broader cohort included in the ELS study.

To be included in the current study, participants were required to have a complete ELS Conversation transcript at both Time 1 and Time 3 (approximately 18 months apart). Because the goal of the study was to evaluate longitudinal consistency in episodic and semantic memory, transcripts also needed to contain the same elicited autobiographical topic at both visits. Transcript versions (e.g., A vs. B; school vs. adult) were documented to ensure accurate topic matching but were not used as exclusion criteria. Only portions of transcripts that reflected clearly identifiable episodic content and were present at both time points were selected for scoring.

After applying inclusion criteria, the final analytic sample consisted of 50 participants, each contributing one matched narrative topic at both time points. These 100 transcripts constituted the full dataset used for episodic and semantic detail coding (AI) and subsequent semantic similarity analyses (SemDis). All data was obtained from the existing ELS study dataset (Abbeduto et al., 2020; Abbeduto et al., 2023), no new data was collected from the participants for this study.

2.3 Procedure

During each visit, participants completed a semi-structured conversation task designed to elicit spontaneous, naturalistic speech in individuals with an intellectual developmental disorder (IDD). Trained examiners followed a flexible script that included open-ended questions and conversational prompts appropriate to the participant's developmental level (e.g., school-age vs. adult versions).

Examiners introduced developmentally appropriate topics (e.g., school, friends, work, hobbies) and used open-ended prompts such as "Tell me about…" to initiate each topic. Follow-up prompts were brief and non-directive (e.g., "What happened next?", "Tell me more", "Who was there?") and were used only to maintain conversational flow. The ELS Conversation Manual emphasizes conversational reciprocity rather than strict standardization, and accordingly, the amount and type of examiner scaffolding varied naturally depending on the participant's responsiveness. These variations reflect the original study's focus on capturing spontaneous, ecologically valid language samples.

All conversations were audio or video recorded and transcribed by trained transcribers using Systematic Analysis of Language Transcripts (SALT) software. Initial transcripts were generated in SALT and then reviewed and corrected to ensure accuracy and adherence to transcription standards. The transcripts used in the current study represent the finalized versions after transcription checks. Audio and video recordings were not available for the present analyses.

2.3.1 Transcript Preparation

Transcript preparation for the current study followed a multi-step, top-down process designed to identify comparable autobiographical narrative material across the two study visits.

2.3.1.1 Version Matching

First, a version-level match was completed to determine whether each participant received the school-age or adult version of the ELS Conversation task at each time point. This ensured that topic opportunities were broadly comparable before examining transcript content in detail.

2.3.1.2 Independent Topic Matching

Second, two trained coders independently reviewed each transcript to identify the topic(s) discussed at each visit and determine whether a common topic(s) appeared at both Time 1 and Time 3 for each participant. Agreement between coders was high and any discrepancies were resolved through discussion.

2.3.1.3 Verification of Topic Matches

Following the independent topic matching, the main coder conducted an additional verification pass while reviewing the transcripts in detail. During this review, the coder confirmed that all matched topics were accurate and that no mismatches or inconsistencies were present. This verification step served as a secondary reliability check and confirmed that the independently matched topics were valid for all participants.

2.3.1.4 Episodic Eligibility Review

Next, the main coder reviewed each transcript to determine if the matched topic segments were prompted to reflect episodic content. Episodic segments were defined as narrative material that refers to a single, time-bound personal event. Segments describing habitual routines or generalized activities were excluded. For example, responses to prompts such as "Tell me everything you do when you get to school" were excluded because they elicit routine or procedural descriptions, whereas responses to "Tell me everything you did at school yesterday" were retained because they reference a discrete event anchored in time.

2.3.1.5 Finalizing Segments for Coding

Only transcript segments that (a) belonged to the independently matched topic, (b) contained episodic material, and (c) were present at both time points were highlighted for AI scoring. Coders had access to the full transcript and read surrounding examiner and participant responses to preserve conversational understanding, including how the examiner prompted the child and any callbacks to earlier topics. However, AI scoring was assigned only to the highlighted participant utterances. Examiner speech was not scored. For the semantic distance analyses, the highlighted episodic segments were later paired with their corresponding examiner prompts for the SemDis input which will be explained in the latter section(s).

2.3.2 Scoring Procedures

2.3.1.1 AI Coding

Narratives were coded following the standard AI scoring system (Levine et al., 2002). Each transcript was segmented into individual details and assigned to one of the internal (event, time, place, perceptual, or emotion/thought) or external (semantic, repetition, or other) categories.

2.3.1.2 Coder Training and Reliability

Coder training followed the procedures outlined in the AI scoring protocol developed by Levine et al. (2002). Training materials were provided by the Levine Lab at The Rotman Research Institute provided by the PI Dr. Brian Levine. Training began with a set of preliminary practice memories, during which each coder compared their segmentation and detail category assignments to a reliable scorer's key and score sheet to familiarize themselves with the coding system. After completing the initial practice phase, coders entered a formal reliability phase consisting of four sets of five memories each (20 memories total). For each set, coders independently scored the memories, and a separate main scorer entered their scores into a correlation spreadsheet containing gold-standard scores established by Levine's reliable coders. The spreadsheet automatically calculated the interrater reliability between the trainee and the gold-standard scorers.

Coders were required to achieve correlations between 0.80 and 0.95 with the gold-standard on each set of five memories. If a coder fell below this threshold, they reviewed and corrected discrepancies, repeated the set until they consistently met the required reliability level across all 20 memories. This multi-stage process ensured that all coders demonstrated accurate and consistent application of the AI coding to scoring the study transcripts.

2.3.1.3 Scoring Workflow

After all coders met reliability criteria, a scoring schedule was created to assign participants and topics across the team. Each participant contributed at least one topic that met inclusion criteria, and some participants had up to three eligible topics across the two time points. The main coder (first author) scored all 50 participants (100 transcripts), and the remaining coders scored additional subsets of transcripts to balance workload and ensure consistency across coders.

Every transcript was double scored by two independent coders. Discrepancies in segmentation or detail categories were reviewed in weekly consensus meetings, and final agreed-upon scores were established for each transcript. These final scores for AI coding were used in all analyses. Although some participants produced more than one topic that met inclusion criteria across visits, only the earliest topic that appeared at both time points was selected for analysis. This ensured that every participant contributed at least one matched topic pair and maintained consistency across the dataset.

2.4 Measures

2.4.1 Autobiographical Interview (AI) Measures

Autobiographical memory performance was coded using the standard AI scoring system (Levine et al., 2002). Consistent with the full AI protocol, coders segmented each transcript into discrete details and assigned each detail to the appropriate internal or external category. Coders also completed the full set of AI global ratings for each transcript, including overall memory quality and time/place-specific ratings, to maintain full fidelity to the standard scoring procedure.

Although the complete AI scoring system was applied, only a subset of AI-derived variables was used in the statistical analyses. The episodic detail memory variable consisted of total internal details, calculated as the sum of all internal (episodic) detail subtypes. The semantic memory variable consisted only of the semantic subset of external details. External detail types such as repetitions, fillers, and other miscellaneous content were coded but not included in the semantic variable used in analyses. Similarly, global AI ratings were generated for each transcript but were not included as analytic variables. This approach preserved the integrity of full AI scoring while aligning analytic variables with study aims. All analytic AI variables reflect consensus-coded scores for the highlighted episodic segments at both time points.

2.4.2 Semantic Distance (SemDis)

Semantic distance was used as an index of narrative coherence. For each transcript, item–response pairs were created by pairing the examiner's initial topic prompt (item) with each of the participant's corresponding episodic utterances (response). After minimal transcript cleaning to remove SALT-specific transcription artifacts, item–response pairs were processed through the archived SemDis platform in R. The archived SemDis implementation used the 840B-token GloVe Common Crawl vectors. The system generated semantic distance values for each pair, which was averaged to create one coherence score per participant per time point.

2.4.3 Cognitive Ability (SB-5)

Cognitive ability was measured using the Stanford-Binet Intelligence Scales, Fifth Edition (SB-5) (Roid & Pomplun, 2012) which were administered as part of the larger ELS study. The FSIQ standard score was used as the primary index of cognitive functioning in the present analyses. FSIQ provides a broad estimate of global intellectual functioning integrating performance across verbal and nonverbal reasoning, knowledge, quantitative reasoning, visual-spatial processing, and working memory. This composite was selected because it captures overall cognitive ability while reducing multicollinearity that would arise from including multiple domain-level composite scores.

The SB-5 demonstrates strong psychometric properties in populations with IDD with internal consistency coefficients and test-retest reliabilities typically above 0.90. Although FSIQ scores were examined descriptively to characterize overall cognitive functioning, they were not included in the final multilevel models. FSIQ did not significantly differ between diagnostic groups and did not account for unique variance in autobiographical memory performance once expressive language scores were included. Additionally, reliance on global IQ composites was avoided given their limited interpretive value for conversational narrative outcomes and their frequent overgeneralization in applied contexts.

Moreover, FSIQ demonstrated substantial conceptual and statistical overlap with expressive language predictors, particularly the CELF FS subtest—which led to multicollinearity and reduced model interpretability. Excluding FSIQ therefore improved model stability and allowed the analyses to focus on predictors most directly related to narrative production.

2.4.4 Expressive Language (CELF-4)

Expressive language skills were assessed using Expressive Vocabulary and FS subtests of the CELF, Fourth Edition (Semel et al., 2003). Raw scores were used because standardized scores were inconsistently available in the original ELS dataset, and discrepancies in administration procedures prevented reliable derivation of normed metrics across participants and time points. Raw scores therefore provided the most complete and interpretable indicators of expressive language ability for the current sample.

2.4.5 Adaptive Communication (VABS-II)

Adaptive communication ability was measured using the Communication Domain Standard Score from the Vineland Adaptive Behavior Scales, Second Edition (VABS-II) (Sparrow et al., 2005). This domain provides a norm-referenced composite reflecting receptive, expressive, and written communication skills used in everyday contexts and is widely used in research involving individuals with IDD due to its strong psychometric properties. In the present study, the VABS-II Communication Domain Standard Score (SS) was reviewed descriptively to characterize participants' adaptive communication abilities, but it was not included as a predictor in the final multilevel models. Descriptive comparisons (see Table 5) indicated that Communication SS were highly similar across diagnostic groups at both time points. At Visit 1, FXS had a mean score of 69.44 (*SD* = 15.71, range = 30–92; *n* = 16) and DS had a mean score of 70.78 (*SD* = 10.26, range = 52–96; *n* = 23). At Visit 3, this pattern remained consistent, with FXS averaging 67.48

(*SD* = 22.29, range = 21–117; *n* = 21) and DS averaging 69.08 (*SD* = 11.54, range = 36–88; *n* = 25). These overlapping means and ranges suggest that adaptive communication abilities did not meaningfully differ between groups and therefore did not require adjustment in the statistical models. Additionally, the VABS-II Communication SS had greater missingness than other predictors, and including it would have reduced the analytic sample size and decreased statistical power. For these reasons, the VABS-II Communication Domain SS was used to describe the sample but was not retained in the multilevel analyses.

2.4.6 Covariates

Several variables were included as covariates based on theoretical and empirical relevance to autobiographical memory performance in IDD. Chronological age at each visit was included to account for developmental differences across the sample, which ranged from 6 to 25 years. Sex was included because males and females can differ in expressive language and cognitive profiles within FXS, due to the diagnostic groups differing in sex distribution. Diagnostic group (DS vs. FXS) was treated as a fixed factor to capture syndrome-specific differences in narrative performance and developmental trajectories.

Child word count from the highlighted episodic narrative segments was included to adjust for individual differences in overall verbal output. Since participants varied widely in the amount of language produced, controlling for word count allowed the analyses to adjust for narrative length independently of memory performance. All covariates were treated as continuous or categorical variables consistent with their measurement properties and were included in the multilevel models as described in the following Analytic Plan.

2.4.7 Preliminary Analyses and Covariate Selection

Prior to multilevel modeling, bivariate Pearson correlations were examined among narrative outcomes and candidate cognitive and language covariates to evaluate shared variance and inform covariate selection. Several language and cognition related measures demonstrated moderate to strong intercorrelations, indicating substantial overlap among predictors. To reduce redundancy and minimize multicollinearity in the multilevel models, covariate selection emphasized parsimony and theoretical relevance, with preference given to measures most directly related to narrative production. Bivariate correlations are presented in Table 11.

2.5 Analytic Plan

2.5.1 Model Specification

Two multilevel models were estimated, one for each primary outcome variable: (a) total internal details derived from the AI coding and (b) semantic distance scores generated using the archived SemDis application. Time (Time 1 and Time 3) was modeled as a within-person predictor to capture change in autobiographical memory performance across the approximately 18-month interval. Between-person predictors included diagnostic group (FXS vs. DS), sex, chronological age, and expressive language ability indexed by the CELF FS raw score. Because variation in the amount of spoken language produced can influence both detail counts and semantic similarity values, child word count for the highlighted narrative segment was included as a covariate in all models to adjust for differences in narrative length.

2.5.2 Multilevel Structure

A MLM framework was selected to account for the nested structure of the data, in which repeated autobiographical memory scores (Level 1) were nested within individuals (Level 2). This approach allowed for the simultaneous estimation of within-person change over time and between-person differences in baseline performance. Time was treated as a fixed effect and centered at Time 1 to facilitate interpretation of intercepts as baseline estimates.

2.5.3 Fixed Effects

Fixed effects were included to evaluate the unique contribution of demographic, cognitive, and linguistic predictors to each outcome. Of particular interest were the effects of diagnostic group and its interaction with time, which tested whether individuals with DS and FXS showed different developmental trajectories in episodic memory detail production and semantic coherence. Chronological age and sex were included to adjust for developmental and demographic influences. Expressive language ability (FS) was included to account for individual differences in linguistic abilities that may support autobiographical memory performance.

2.5.4 Random Effects

All models included a random intercept to account for individual differences in baseline levels of memory performance. Random slopes for time were tested to determine whether allowing individuals to vary in their rate of change improved the model fit. A random slope was only retained when its inclusion significantly improved model performance based on model comparison criteria. When random slopes resulted in convergence failures or did not significantly improve fit, the model was estimated with a random intercept only. Random slopes for time were tested but were not retained in the final models because they did not improve model fit and occasionally produced convergence issues.

2.5.5 Covariates

Child word count for the highlighted topic segment was included as a covariate in each model. Word count adjusts for differences in narrative length, which is critical for studies using Levine AI scoring because individuals who speak more, may artificially appear to produce more details. This adjustment is likewise essential for semantic distance analyses, as semantic similarity values can shift based on the quantity and distribution of lexical content.

2.5.6 Model Fitting and Diagnostics

Predictors were entered simultaneously in all multilevel models to estimate their unique associations with the outcome variables, and all data were screened before fitting the final models. During these preliminary checks, including visual inspection of Q–Q plots and distributions, one participant stood out due to an unusually large discrepancy in their CELF-FS score across visits. The participant received a score of 5 at Visit 1 and 56 at Visit 3, and this extreme shift created a noticeable distortion in the residual pattern during early diagnostics. Because of the magnitude of this change, the participant's broader expressive and adaptive profile was reviewed. The very low CELF-FS T1 score did not align with the participant's Vineland Communication Domain standard scores, their age, or the quality of their narrative responses, which included one of the highest internal detail counts at T3. Together, these indicators suggested that the T1 CELF-FS score was likely the result of a scoring or administration error rather than a true reflection of expressive language ability. To avoid removing the participant entirely, and to prevent this singular irregular score from disproportionately influencing model performance, the T1 CELF-FS score was replaced with the sample mean for that timepoint (29). Given the minimal amount of missing or inconsistent data in the dataset and the clear evidence supporting an error in the original score, this targeted imputation was considered appropriate before fitting the models.

After this correction, the multilevel models were estimated using all available data. Model fit was evaluated using the −2log Likelihood, the Akaike Information Criterion (AIC), and the Bayesian Information Criterion (BIC). Standard diagnostic procedures were used to examine linearity, normality of residuals, and homoscedasticity. Due to SPSS MIXED not providing a multilevel equivalent of Cook's Distance, the influence was assessed by examining standardized Level-1 residuals aggregated for each participant, which is an established approach for identifying cases that may unreasonably affect parameter estimates. In the present study, absolute standardized residuals ranged from 0.14 to 2.83, which is below the commonly used $|z| > 3$ threshold for identifying potentially influential observations. These results indicated that no participants exerted unjustified influence on model estimates, supporting the stability and robustness of the final models.

2.5.7 Software and Statistical Thresholds

All multilevel models were conducted in SPSS Version 31. Semantic distance values were computed in R using the archived SemDis platform, which was run locally following the developers' installation instructions. Statistical significance was evaluated at an alpha level of 0.05, and effect sizes and confidence intervals were reported where appropriate.

Table 6

Descriptive Statistics for VABS-II Communication Domain Standard Scores by Diagnostic Group and Timepoint

Diagnosis	timepoint	Mean	Std. Deviation	Minimum	Maximum	N
Fragile X Syndrome (FXS)	1	69.44	15.714	30	92	16
	3	67.48	22.290	21	117	21
	Total	68.32	19.491	21	117	37
Down Syndrome (DS)	1	70.78	10.264	52	96	23
	3	69.08	11.539	36	88	25
	Total	69.90	10.865	36	96	48
Total	1	70.23	12.606	30	96	39
	3	68.35	17.103	21	117	46
	Total	69.21	15.149	21	117	85

Note. Means and standard deviations reflect Vineland Adaptive Behavioral Scale, Second Edition (VABS-II) Communication Domain Standard Scores (SS) at each timepoint (1, 3) by diagnostic group (FXS, DS). Higher scores indicate stronger adaptive communication abilities.

Chapter 3. RESULTS

Two multilevel models were estimated to examine predictors of autobiographical memory performance across the two study visits. Participants (*N* = 50) contributed repeated observations at Time 1 and Time 3, and analyses were conducted using the SPSS Mixed Models procedure. Descriptive statistics for all narrative outcomes are shown below in Table 7.

Table 7

Descriptive Statistics for Narrative Outcomes (N = 50)

Outcomes	Time 1			Time 3		
	M (*SD*)	Min	Max	M (*SD*)	Min	Max
Episodic Details (Internal AI)	6.44 (5.33)	0.00	23.00	8.90 (7.85)	0.00	32.00
Semantic Details (External AI)	0.24 (0.56)	0.00	2.00	0.54 (1.40)	0.00	9.00
Semantic Coherence (SemDis)	1.03 (0.34)	0.12	1.85	0.95 (0.33)	0.19	1.95
Word Count	24.36 (19.65)	3.00	96.00	41.36 (37.68)	2.00	141.00

3.1 Predictors of Episodic (Internal) Details

Aim 1 examined whether diagnostic group, expressive language ability, narrative length, and timepoint predicted the number of internal episodic details children produced during the autobiographical memory task. Because internal detail counts were expected to vary both within and between participants, a multilevel model was used with repeated observations (Time 1 and Time 3) nested within individuals. A random intercept was included to account for individual variability, and time was modeled as a within-person factor. The analytic approach involved first evaluating whether the raw outcome met model assumptions, then determining whether transformations were required, and finally interpreting the fixed effects from the best-fitting model.

Consistent with recommended analytic procedures, the multilevel model was first estimated using the untransformed internal detail variable to evaluate whether the raw outcome met the assumptions for normality and homoscedasticity. Inspection of residuals indicated deviations from normality and heteroscedasticity in the untransformed internal detail variable. Due to these diagnostics indicating violations of normality and constant variance, square root and natural log transformations were evaluated.

The natural log transformation noticeably improved the distribution of internal details. Model fit was substantially better (AIC = 254.14), and the normal Q–Q plot (see Figure 7) showed residuals closely following the diagonal with only slight tail deviation. The detrended Q–Q plot (see Figure 8) no longer displayed the curved pattern seen in earlier models, and the residuals-versus-predicted plot (see Figure 9) showed an even horizontal spread rather than bunching. The histogram of residuals showed a SD of approximately 0.70 (see Figure 10). Overall, these checks indicated that the natural log transformed model met normality and variance assumptions so this was the selected model moving forward. A random intercept was included to account for individual differences across participants, and the unstructured covariance estimates indicated moderate within-person variability across time. Model fit indices (AIC = 254.14) supported the ln-transformed model as the best fitting for Aim 1.

The final natural log transformed model included both child word count (c_wd_count) and semantic external details (semantic_ext). Fixed effects results from the final ln-transformed model are presented in Table 9. Child word count (c_wd_count) was the only significant predictor of internal episodic details, $F(1, 78.21) = 19.69$, $p < .001$, suggesting that children who produced more words also produced more internal details (Estimate = .012, *SE* = .003). In contrast, semantic external details were not a significant predictor, $F(1, 67.44) = 0.99$, $p = .325$, indicating that once narrative length was accounted for, semantic elaboration did not explain additional variance in internal detail production. Diagnostic group ($F(1, 53.95) = 1.42$, $p = .239$), timepoint ($F(1, 49.19) = 0.05$, $p = .833$), and the diagnosis × timepoint interaction ($F(1, 48.26) = 0.15$, $p = .696$) were all nonsignificant. Gender was also nonsignificant, $F(1, 45.09) = 0.72$, $p = .399$, as was expressive language (CELF-FS), $F(1, 70.80) = 1.46$, $p = .232$, and the diagnosis × CELF-FS interaction, $F(1, 63.42) = 0.21$, $p = .650$. Together, these results indicate that, after adjusting for narrative length, none of the demographic or language variables uniquely predicted internal episodic detail production.

Taken together, these findings indicate that internal episodic detail production was strongly driven by narrative length rather than diagnostic group, timepoint, or expressive language skills. Once word count was included, neither semantic external details nor language ability (FS) uniquely predicted internal details. Overall, internal detail production appeared stable across the two time points and comparable across groups.

The covariance structure indicated substantial within-person variability in internal detail production across the two visits (see Table 8). Variance estimates at Time 1 (.599) and Time 3 (.475) were both significant, reflecting meaningful variability at each assessment. In contrast, the estimated covariance between timepoints was small and nonsignificant (.091, $p = .273$), suggesting weak stability in internal detail production across the 18-month interval.

Table 8

Estimates of Covariance Parameters[a]

Parameter		Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Repeated Measures	UN (1,1)	.599	.126	4.772	<.001	.398	.904
	UN (2,1)	.091	.083	1.096	.273	-.071	.253
	UN (2,2)	.475	.100	4.733	<.001	.314	.719

Note. UN (1,1) and UN (2,2) represent the estimated residual variances for Time 1 and Time 3, respectively, from the unstructured covariance matrix. UN (2,1) represents the estimated covariance between the two timepoints. Larger variance values indicate greater within-person variability at each visit, while the covariance reflects the degree to which individuals' scores at Time 1 predict their scores at Time 3.
a. Dependent Variable: ln_total_int.

To supplement the multilevel model and to better understand change over time, a within-person change score was calculated (T3 - T1). The mean change in internal details was small (*M* = 2.46), indicating no pervasive increase or decrease across the sample, which aligns with the nonsignificant timepoint effect in the multilevel model. However, the standard deviation was large (*SD* = 9.00), and scores ranged from –17 to +28, reflecting substantial variability across participants. Some individuals showed notable decreases while others showed large increases across the two time points. This pattern suggests that, although group-level means were stable over time, individual trajectories varied widely, supporting the interpretation that internal detail production is influenced more by differences in narrative output than by diagnosis or time.

3.2 Predictors of Semantic Coherence (SemDis)

Aim 2 evaluated whether diagnostic group, expressive language ability, narrative length, or time predicted semantic coherence (SemDis). A multilevel model was fit with SemDis as the dependent variable, using timepoint (T1 and T3) as a repeated factor nested within participants and including a random intercept to account for individual differences. An unstructured covariance structure was used.

Residual diagnostics indicated that model assumptions were reasonably satisfied, with no evidence of substantial non-normality or heteroscedasticity (see Figures 11–14). Together, these patterns supported normality and homoscedasticity without requiring transformation. Residual variances were small and nearly identical across visits (T1 = .112; T3 = .100), and the covariance between timepoints was close to zero (−.003). This pattern suggests that semantic coherence did not change much from Time 1 to Time 3.

Fixed-effects results for Semantic Distance are shown in Table 12. Diagnostic group was not a significant predictor of coherence, $F(1, 52.55) = 3.81$, $p = .056$. Coherence also did not significantly change from Time 1 to Time 3, $F(1, 50.11) = 2.36$, $p = .131$. Word count was not significant, $F(1, 76.42) = 3.06$, $p = .084$. Gender was not significant, $F(1, 44.88) = 1.84$, $p = .182$. The main effect of CELF–FS was also not significant, $F(1, 65.64) = 1.27$, $p = .263$.

The only significant finding in this model was the interaction between diagnostic group and FS, $F(1, 61.26) = 5.55$, $p = .022$. When the groups were examined separately, children with DS showed lower Semantic Distance values (more coherent narratives) as their FS scores increased. Children with FXS showed a slight increase in Semantic Distance (less coherent narratives) as their FS scores increased.

Overall, semantic coherence was relatively stable across time and largely similar between diagnostic groups. However, expressive language ability showed a group specific association with SemDis, stronger expressive language skills predicted higher semantic coherence among individuals with DS but not for individuals with FXS. These findings suggest that the role of expressive language in supporting semantic organization may differ across neurodevelopmental conditions.

Chapter 4. DISCUSSION

The present study explored autobiographical narrative performance in children with DS and FXS using conversational samples collected through Abbeduto et al.'s (2020, 2023) ELS task. Two aspects of narrative accounts were examined: episodic detail production and narrative coherence measured through Semantic Distance. Although prior literature highlights differences in language and cognitive functioning between DS and FXS, the results of this study revealed a more nuanced picture.

Prior work has shown that DS and FXS are associated with distinct language profiles, particularly in expressive grammatical ability, yet substantial variability exists within each diagnostic group. For example, Finestack, Sterling, and Abbeduto (2013) reported that children with DS and FXS differed in grammatical skills but also demonstrated considerable within-group variability in conversational language. Similarly, Martzoukou, Nousia, and Marinis (2020) found that narrative performance in individuals with DS was strongly influenced by individual differences in expressive morphosyntax rather than by diagnostic classification alone. The findings of the present study align with this work. Across both aims, children demonstrated notable variability in their narrative output, and this variability appeared to play a more influential role in shaping autobiographical performance than diagnostic group.

4.1 Episodic (Internal) Details

The first aim focused on episodic (internal) detail production. It was anticipated that children with FXS would produce more internal details than those with DS, but this expectation was not supported. Children across both groups produced a similar number of episodic details, and there was no meaningful change between Time 1 and Time 3. Word count was the only significant predictor of internal details for Aim 1. Children who produced more words also tended to provide more episodic (internal) content. Given the large individual differences observed in the amount of language produced, the association between word count and episodic details suggests that expressive output ability played an important role in how much episodic details that the information children shared during their narratives.

This pattern is consistent with the intended function of the ELS Conversation task as a naturalistic measure of expressive language. The task is designed to elicit spontaneous, conversational speech that reflects how children respond in everyday interactions. Within this structure, the amount of episodic information revealed by each child may depend largely on how easily they enter the conversational exchange, how comfortable they feel sustaining it, and how much expressive access they have at that given moment, much like they might in some aspects of everyday memory recall, which may or may not translate to reporting in a forensic context. Some children may recall more episodic content than they verbally express, particularly if expressive language demands or attentional fluctuations may limit the amount of speech they are able to produce. In this context, the findings point toward an association between verbal engagement and episodic detail production rather than a direct reflection of underlying memory differences.

Of interest is that external details did not relate to semantic skill in the CELF formulated sentences or the production of semantic details, but the word count produced. This finding suggests that the elicitation of speech was the single biggest predictor of episodic details, not the organization of that speech. Individuals with language disorders are often judged by speech organization, and our finding here suggests no relation.

4.2 Narrative Coherence (SemDis)

The second aim examined narrative coherence. Based on previous findings, it was hypothesized that children with DS would show lower coherence than those with FXS, but the results did not support this prediction. Coherence did not reliably differ by group, and there was no change over time. Word count, gender, and CELF-FS by themselves did not predict coherence. The only significant effect observed was the interaction between diagnostic group and FS (syntax generation) scores. Among children with DS, higher FS scores were associated with lower Semantic Distance values, which correspond to more coherent narratives. In contrast, children with FXS exhibited a slight increase in Semantic Distance as FS scores increased, indicating lower coherence.

The opposite directions of these associations suggest that FS abilities may relate differently to narrative organization across groups. Children with DS often produced brief, literal responses that stayed closely connected to the conversational prompt, and higher FS performance appeared to be associated with clearer and more coherent output for these children. Children with FXS, on the other hand, frequently produced more complex utterances that sometimes included repetition, tangential language, or shifts in attention, aspects of their language previously reported in the literature (Hoffmann et al., 2022; Abbeduto et al., 2007) SemDis, which captures semantic similarity between consecutive utterances, may be sensitive to these pragmatic and attentional differences. Because of this, higher FS ability in the FXS group may have coincided with narrative patterns that increased semantic distance rather than reducing it. These associations should not be interpreted as causal but instead reflecting patterns within the sample that appear consistent with known differences in communication profiles.

Across both aims, the results indicate that narrative performance during the ELS Conversation task is closely tied to language production and engagement. Notably, once children began talking and sustained the interaction, they were generally more likely to provide additional episodic information. This pattern was particularly noticeable among children with DS, who did not exhibit the same level of echolalia or perseverative speech sometimes seen in FXS. When children with DS became verbally engaged, they often remained on topic and contributed many of their accessible narrative details. In contrast, several children with FXS produced longer stretches of speech that included repetition or brief shifts in focus. When this happened, the SemDis model had more linguistic material to compare against the examiner's prompt, which may have increased semantic distance and made the narrative appear less coherent. These patterns suggest that the degree to which children stayed verbally engaged with the topic was associated with both the amount of information they expressed and how closely their responses aligned with the conversational prompt.

A related point concerns the variability within each diagnostic group. Children varied widely in their expressive output and narrative coherence, and these differences were often larger within groups than between them. Although evaluating representativeness was not an explicit aim of this study, the distribution of abilities in the analytic subsample appeared consistent with the variability reported in the broader parent ELS dataset. This similarity suggests that the range of expressive and narrative abilities observed in this study likely reflects the natural heterogeneity of DS and FXS, rather than sampling bias or cohort effects. Taken together, the findings highlight the importance of considering individual expressive language profiles when interpreting narrative performance.

4.3 Implications

4.3.1 Clinical Implications

The clinical implications of these observations are considerable. Due to children sharing more narrative information when they felt comfortable engaging in conversation, clinicians may benefit from allocating time to rapport building before eliciting autobiographical material. Supporting children through predictable routines, warm-up conversation, or play-based interaction may help them enter the verbal exchange more easily. Once they begin speaking, providing additional wait time, simplifying prompts, or using visual aids may further support expressive access. Children who have difficulty relying solely on spoken language may benefit from alternative communication strategies, such as drawing, sequencing cards, Augmentative and Alternative Communication devices, or through gesture (Danby & Sharman, 2024; Iacono et al., 2016; Korkman et al., 2025). These approaches may offer a clearer window into children's autobiographical knowledge and emotional experiences, especially when expressive limitations constrain what they can share verbally.

4.3.2 Forensic Implications

In forensic and investigative contexts, the results suggest that a child's narrative performance during an interview may depend as much on expressive and attentional factors as on memory itself. Rapport-building practices that help children feel comfortable and confident may play a key role in eliciting more complete accounts. Interviewers who provide adequate wait time, reduce linguistic complexity, offer visual supports, and adjust the pace of questioning may create conditions that allow children to express what they remember more effectively (Witwer et al., 2025). Such accommodations may be especially important when interviewing children with DS or FXS, whose expressive profiles can vary widely and may otherwise lead to underestimation of their recall abilities.

4.3.3 Research Implications

From a research perspective, the high degree of within-group variability suggests that relying solely on diagnostic group comparisons may obscure meaningful individual differences. Narrative performance in a conversational setting likely reflects a combination of language ability, attention, and personal communication patterns, which can differ substantially across children regardless of diagnosis. Future work may benefit from designs that account for these individual differences by incorporating measures of attention or working memory into analytic models. Since conversational tasks require children to manage both language and memory at the same time, adding more structured memory measures could help clarify which aspects of their performance reflect expressive limitations and which reflect memory ability.

4.4 Study Strengths and Limitations

The study has several strengths, including its use of ecologically valid conversational data, its integration of human-coded and computational narrative measures, and its focus on two neurodevelopmental conditions rarely examined together in the autobiographical memory literature. Nonetheless, several limitations warrant consideration. The ELS task centers on expressive language rather than memory retrieval, and examiner pacing varied across interactions. The absence of audio or video limited the ability to verify engagement or clarify unintelligible speech. Unclear segments reduced the amount of analyzable material, and the modest sample size constrained the detection of subtle effects. Even so, the alignment of this sample's variability with the broader ELS dataset provides reassurance that the patterns observed here are grounded in typical expressive diversity across these groups. Further, the sample selected for this project included individuals able to produce 3-word phrases, so these findings are not representative of the whole range of ability in the syndromes but a segment of the population with some verbal ability.

4.5 Future Directions

Future research may benefit from pairing conversational samples with more structured memory tasks and incorporating multimodal communication supports Another promising direction involves making fuller use of rating systems that capture the richness and quality of narrative details, including distinctions between episodic information, external content, and repetitions. Applying these kinds of scoring approaches may help to clarify the extent to which children are attempting to communicate meaningful autobiographical information that may not be fully visible within conversational transcripts alone. In addition, continued development of natural language processing tools that are trained on neurodivergent or non-neuronormative speech may allow for more accurate measurement of narrative features in DS and FXS, as current models are largely trained on neurotypical language. Integrating these methods in future work may support a more precise understanding of how different language and cognitive processes relate to the autobiographical information children are able to share in naturalistic conversation.

4.6 General Summary

In summary, this study suggests that autobiographical narrative performance in DS and FXS reflects a complex interplay of expressive language, engagement, and individual communication style. The results illustrate substantial overlap and variability across groups, highlighting the importance of individualized approaches when evaluating narrative abilities in children with neurodevelopmental disorders. Although the findings do not indicate robust group differences, they underscore the value of creating supportive conversational contexts that allow children to express what they know and remember in ways that align with their strengths.

References

Abbeduto, L., Berry-Kravis, E., Sterling, A., Sherman, S., Edgin, J. O., McDuffie, A., Hoffmann, A., Hamilton, D., Nelson, M., Aschkenasy, J., & Thurman, A. J. (2020). Expressive language sampling as a source of outcome measures for treatment studies in fragile X syndrome: Feasibility, practice effects, test-retest reliability, and construct validity. *Journal of Neurodevelopmental Disorders*, *12*(1), 10. https://doi.org/10.1186/s11689-020-09313-6

Abbeduto, L., Brady, N., & Kover, S. T. (2007). Language development and fragile X syndrome: Profiles, syndrome-specificity, and within-syndrome differences. *Mental Retardation and Developmental Disabilities Research Reviews*, *13*(1), 36–46. https://doi.org/10.1002/mrdd.20142

Abbeduto, L., Del Hoyo Soriano, L., Berry-Kravis, E., Sterling, A., Edgin, J. O., Abdelnur, N., Drayton, A., Hoffmann, A., Hamilton, D., Harvey, D. J., & Thurman, A. J. (2023). Expressive language sampling and outcome measures for treatment trials in fragile X and down syndromes: Composite scores and psychometric properties. *Scientific Reports*, *13*(1), 9267. https://doi.org/10.1038/s41598-023-36087-3

Acevedo-Molina, M. C., Matijevic, S., & Grilli, M. D. (2020). Beyond episodic remembering: Elaborative retrieval of lifetime periods in young and older adults. *Memory (Hove, England)*, *28*(1), 83–93. https://doi.org/10.1080/09658211.2019.1686152

Allen, E. G., Sherman, S., Abramowitz, A., Leslie, M., Novak, G., Rusin, M., Scott, E., & Letz, R. (2005). Examination of the Effect of the Polymorphic CGG Repeat in the FMR1 Gene on Cognitive Performance. *Behavior Genetics*, *35*(4), 435–445. https://doi.org/10.1007/s10519-005-2792-4

Bailey, D. B., Raspa, M., Bishop, E., Mitra, D., Martin, S., Wheeler, A., & Sacco, P. (2012). Health and economic consequences of Fragile X syndrome for caregivers. *Journal of Developmental & Behavioral Pediatrics*, *33*(9), 705–712. https://doi.org/10.1097/DBP.0b013e318272dcbc

Bailey, D. B., Raspa, M., & Olmsted, M. G. (2010). Using a parent survey to advance knowledge about the nature and consequences of Fragile X syndrome. *American Journal on Intellectual and Developmental Disabilities*, *115*(6), 447–460. https://doi.org/10.1352/1944-7558-115.6.447

Beaty, R. E., & Johnson, D. R. (2021). Automating creativity assessment with SemDis: An open platform for computing semantic distance. *Behavior Research Methods*, *53*(2), 757–780. https://doi.org/10.3758/s13428-020-01453-w

Berglund, E., Eriksson, M., & Johansson, I. (2001). Parental reports of spoken language skills in children with Down syndrome. *Journal of Speech, Language, and Hearing Research*, *44*(1), 179–191. https://doi.org/10.1044/1092-4388(2001/016)

Berry-Kravis, E., Doll, E., Sterling, A., Kover, S. T., Schroeder, S. M., Mathur, S., & Abbeduto, L. (2013). Development of an expressive language sampling procedure in Fragile X syndrome: a pilot study. *Journal of Developmental & Behavioral Pediatrics*, *34*(4), 245–251. https://doi.org/10.1097/DBP.0b013e31828742fc

Bertelli, M. O., Munir, K., Harris, J., & Salvador-Carulla, L. (2016). "Intellectual developmental disorders": Reflections on the international consensus document for redefining "mental retardation-intellectual disability" in ICD-11. *Advances in Mental Health and Intellectual Disabilities*, *10*(1), 36–58. https://doi.org/10.1108/AMHID-10-2015-0050

Boudreau, D. M., & Chapman, R. S. (2000). The relationship between event representation and linguistic skill in narratives of children and adolescents with Down syndrome. *Journal of Speech,* *Language, and Hearing Research: JSLHR*, *43*(5), 1146–1159. https://doi.org/10.1044/jslhr.4305.1146

Bull, M. J., Trotter, T., Santoro, S. L., Christensen, C., Grout, R. W., & The Council of Genetics. (2022). Health supervision for children and adolescents with Down syndrome. *Pediatrics*, *149*(5), e2022057010. https://doi.org/10.1542/peds.2022-057010

Chapman, R. S., & Hesketh, L. J. (2000). Behavioral phenotype of individuals with Down syndrome. *Mental Retardation and Developmental Disabilities Research Reviews*, *6*(2), 84–95. https://doi.org/10.1002/1098-2779(2000)6:2<84::AID-MRDD2>3.0.CO;2-P

Danby, M. C., & Sharman, S. J. (2024). An exploration of interview breaks in child forensic interviews. *Policing: A Journal of Policy and Practice*, *18*, paae081. https://doi.org/10.1093/police/paae081

Finestack, L. H., Palmer, M., & Abbeduto, L. (2012). Macrostructural narrative language of adolescents and young adults with Down syndrome or Fragile X syndrome. *American Journal of Speech-Language Pathology*, *21*(1), 29–46. https://doi.org/10.1044/1058-0360(2011/10-0095)

Gioia, G., Isquith, P. K., Guy, S. C., & Kenwiorthy, L. (2015). *BRIEF 2*. Psychological Assessment Resources.

Grilli, M. D., & Verfaellie, M. (2015). Supporting the self-concept with memory: Insight from amnesia. *Social Cognitive and Affective Neuroscience*, *10*(12), 1684–1692. https://doi.org/10.1093/scan/nsv056

Grilli, M. D., Wank, A. A., Bercel, J. J., & Ryan, L. (2018). Evidence for reduced autobiographical memory episodic specificity in cognitively normal middle-aged and older individuals at increased risk for Alzheimer's disease dementia. *Journal of the International Neuropsychological Society*, *24*(10), 1073–1083. https://doi.org/10.1017/S1355617718000577

Gross, C., Yao, X., Pong, D. L., Jeromin, A., & Bassell, G. J. (2011). Fragile X mental retardation protein regulates protein expression and mrna translation of the potassium channel Kv4.2. *The Journal of Neuroscience*, *31*(15), 5693–5698. https://doi.org/10.1523/JNEUROSCI.6661-10.2011

Hoffmann, A., Thurman, A. J., Sterling, A., Kover, S. T., Finestack, L., Berry-Kravis, E., Edgin, J. O., Drayton, A., Fombonne, E., & Abbeduto, L. (2022). Analysis of a repetitive language coding system: comparisons between Fragile X syndrome, Autism, and Down syndrome. *Brain Sciences*, *12*(5), 575. https://doi.org/10.3390/brainsci12050575

Hunter, J., Rivero-Arias, O., Angelov, A., Kim, E., Fotheringham, I., & Leal, J. (2014). Epidemiology of fragile X syndrome: A systematic review and meta-analysis. *American Journal of Medical Genetics Part A*, *164*(7), 1648–1658. https://doi.org/10.1002/ajmg.a.36511

Iacono, T., Trembath, D., & Erickson, S. (2016). The role of augmentative and alternative communication for children with autism: Current status and future trends. *Neuropsychiatric Disease and Treatment*, *12*, 2349–2361. https://doi.org/10.2147/NDT.S95967

IBM SPSS Statistics for Windows (Version 28.0). (2021). [Computer software]. Armonk, NY: IBM Corp.

Irazábal, M., Marsà, F., García, M., Gutiérrez-Recacha, P., Martorell, A., Salvador-Carulla, L., & Ochoa, S. (2012). Family burden related to clinical and functional variables of people with intellectual disability with and without a mental disorder. *Research in Developmental Disabilities*, *33*(3), 796–803. https://doi.org/10.1016/j.ridd.2011.12.002

Jarrold, C., Phillips, C., & Baddeley, A. (2002). *Verbal Short-Term Memory in Down Syndrome*. https://doi.org/10.1044/1092-4388(2002/042)

Korkman, J., Otgaar, H., Geven, L. M., Bull, R., Cyr, M., Hershkowitz, I., Mäkelä, J.-M., Mattison, M., Milne, R., Santtila, P., van Koppen, P., Memon, A., Danby, M., Filipovic, L., Garcia, F. J., Gewehr, E., Gomes Bell, O., Järvilehto, L., Kask, K., … Volbert, R. (2025a). White paper on forensic child interviewing: Research-based recommendations by the European Association of Psychology and Law. *Psychology, Crime & Law*, *31*(8), 987–1030. https://doi.org/10.1080/1068316X.2024.2324098

Korkman, J., Otgaar, H., Geven, L. M., Bull, R., Cyr, M., Hershkowitz, I., Mäkelä, J.-M., Mattison, M., Milne, R., Santtila, P., van Koppen, P., Memon, A., Danby, M., Filipovic, L., Garcia, F. J., Gewehr, E., Gomes Bell, O., Järvilehto, L., Kask, K., … Volbert, R. (2025b). White paper on forensic child interviewing: Research-based recommendations by the European Association of Psychology and Law. *Psychology, Crime & Law*, *31*(8), 987–1030. https://doi.org/10.1080/1068316X.2024.2324098

Lee, N. R., Adeyemi, E. I., Lin, A., Clasen, L. S., Lalonde, F. M., Condon, E., Driver, D. I., Shaw, P., Gogtay, N., Raznahan, A., & Giedd, J. N. (2016). Dissociations in cortical morphometry in youth with Down syndrome: evidence for reduced surface area but increased thickness. *Cerebral Cortex*, *26*(7), 2982–2990. https://doi.org/10.1093/cercor/bhv107

Levine, B., Svoboda, E., Hay, J. F., Winocur, G., & Moscovitch, M. (2002). Aging and autobiographical memory: Dissociating episodic from semantic retrieval. *Psychology and Aging*, *17*(4), 677–689.

Mai, C. T., Isenburg, J. L., Canfield, M. A., Meyer, R. E., Correa, A., Alverson, C. J., Lupo, P. J., Riehle-Colarusso, T., Cho, S. J., Aggarwal, D., Kirby, R. S., & National Birth Defects Prevention Network. (2019). National population-based estimates for major birth defects, 2010-2014. *Birth Defects Research*, *111*(18), 1420–1435. https://doi.org/10.1002/bdr2.1589

Martorell, A., Gutiérrez-Recacha, P., Irazábal, M., Marsà, F., & García, M. (2011). Family impact in intellectual disability, severe mental health disorders and mental health disorders in ID. A comparison. *Research in Developmental Disabilities*, *32*(6), 2847–2852. https://doi.org/10.1016/j.ridd.2011.05.021

Martzoukou, M., Nousia, A., & Marinis, T. (2020). Narrative abilities of adults' with Down syndrome as a window to their morphosyntactic, socio-cognitive, and prosodic abilities. *Frontiers in Psychology*, *11*. https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2020.02060

McDonnell, C. G., Boan, A. D., Bradley, C. C., Seay, K. D., Charles, J. M., & Carpenter, L. A. (2019). Child maltreatment in autism spectrum disorder and intellectual disability: Results from a population-based sample. *Journal of Child Psychology and Psychiatry*, *60*(5), 576–584. https://doi.org/10.1111/jcpp.12993

McGrath, R. J., Stransky, M. L., Cooley, W. C., & Moeschler, J. B. (2011). National profile of children with down syndrome: disease burden, access to care, and family impact. *The Journal of Pediatrics*, *159*(4), 535-540.e2. https://doi.org/10.1016/j.jpeds.2011.04.019

McLennan, Y., Polussa, J., Tassone, F., & Hagerman, R. (2011). Fragile X Syndrome. *Current Genomics*, *12*(3), 216–224. https://doi.org/10.2174/138920211795677886

Miles, S., & Chapman, R. S. (2002). Narrative content as described by individuals with Down syndrome and typically developing children. *Journal of Speech, Language, and Hearing Research*, *45*(1), 175–189. https://doi.org/10.1044/1092-4388(2002/013)

Murphy, G., Loftus, E. F., Grady, R. H., Levine, L. J., & Greene, C. M. (2019). False memories for fake news during Ireland's abortion referendum. *Psychological Science*, *30*(10), 1449–1459. https://doi.org/10.1177/0956797619864887

Nadel, L. (2003). Down's syndrome: A genetic disorder in biobehavioral perspective. *Genes, Brain and Behavior*, *2*(3), 156–166. https://doi.org/10.1034/j.1601-183X.2003.00026.x

Peugh, J. L. (2010). A practical guide to multilevel modeling. *Journal of School Psychology*, *48*(1), 85–112. https://doi.org/10.1016/j.jsp.2009.09.002

R: A language and environment for statistical computing. (2023). [Computer software]. R Foundation for Statistical Computing, Vienna, Austria.

Roid, G. H., & Pomplun, M. (2012). *The Stanford-Binet Intelligence Scales, Fifth Edition*.

Semel, E. H., & Secord, W. E. (2003). *Clinical evaluation of language fundamentals, fourth edition (CELF-4*. The Psychological Corporation/A Harcourt Assessment Company in Toronto, Canada.

Simpson, R., Oyekan, A. A., Ehsan, Z., & Ingram, D. G. (2018). Obstructive sleep apnea in patients with Down syndrome: Current perspectives. *Nature and Science of Sleep*, 287–293. https://doi.org/10.2147/NSS.S154723

Sparrow, S. S., Cicchetti, D., & Balla, D. A. (2005). *Vineland Adaptive Behavior Scales, Second Edition (Vineland-II)*. PsycTESTS.

Thomsen, D. K. (2015). Autobiographical periods: a review and central components of a theory. *Review of General Psychology*, *19*(3), 294–310. https://doi.org/10.1037/gpr0000043

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 382–403). New York: Academic Press.

Tulving, E. (1985). How many memory systems are there? *American Psychologist*, *40*(4), 385–398. https://doi.org/10.1037/0003-066X.40.4.385

Wang, M.-T., L. Degol, J., Amemiya, J., Parr, A., & Guo, J. (2020). Classroom climate and children's academic and psychological wellbeing: A systematic review and meta-analysis. *Developmental Review*, *57*, 100912.

Appendix

Table 9

Fixed Effects Estimates From the Multilevel Model Predicting Internal Details[a]

						95% Confidence Interval	
Parameter	Estimate	Std. Error	df	t	Sig.	Lower Bound	Upper Bound
Intercept	1.829	.212	59.939	8.646	<.001	1.406	2.252
[diagnosis_group=1]	-.352	.335	61.476	-1.050	.298	-1.021	.318
[diagnosis_group=2]	0[b]	0	.	.	.	.	.
semantic_ext	.173	.078	64.586	2.221	.030	.017	.329
[timepoint=0]	-.302	.200	46.710	-1.511	.137	-.703	.100
[timepoint=1]	0[b]	0	.	.	.	.	.
[timepoint=0] * [diagnosis_group=1]	.309	.295	47.499	1.045	.301	-.285	.903
[timepoint=1] * [diagnosis_group=1]	0[b]	0	.	.	.	.	.
[timepoint=0] * [diagnosis_group=2]	0[b]	0	.	.	.	.	.
[timepoint=1] * [diagnosis_group=2]	0[b]	0	.	.	.	.	.
[els_gender=0]	-.163	.217	45.714	-.748	.458	-.600	.275
[els_gender=1]	0[b]	0	.	.	.	.	.
celf_fs	.033	.013	60.969	2.564	.013	.007	.058
[diagnosis_group=1] * celf_fs	-.022	.016	60.050	-1.382	.172	-.055	.010
[diagnosis_group=2] * celf_fs	0[b]	0	.	.	.	.	.

Note. Fixed effects estimates are shown for predictors of natural log-transformed internal details. Rows marked 0[b] indicate the comparison categories for group and timepoint variables.

a. Dependent Variable: ln_total_int.

b. This parameter is set to zero because it is redundant.

Table 10

Estimates of Fixed Effects From the Multilevel Model Predicting Semantic Distance[a]

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Intercept	.931	.087	59.359	10.639	<.001	.756	1.106
[diagnosis_group=1]	-.158	.131	63.094	-1.209	.231	-.420	.103
[diagnosis_group=2]	0[b]	0	.	.	.	.	.
[timepoint=0]	.161	.093	50.571	1.736	.089	-.025	.348
[timepoint=1]	0[b]	0	.	.	.	.	.
c_wd_count	.002	.001	76.418	1.750	.084	.000	.004
[timepoint=0] * [diagnosis_group=1]	-.113	.135	48.947	-.842	.404	-.384	.157
[timepoint=1] * [diagnosis_group=1]	0[b]	0	.	.	.	.	.
[timepoint=0] * [diagnosis_group=2]	0[b]	0	.	.	.	.	.
[timepoint=1] * [diagnosis_group=2]	0[b]	0	.	.	.	.	.
[els_gender=0]	.109	.080	44.880	1.357	.182	-.053	.270
[els_gender=1]	0[b]	0	.	.	.	.	.
celf_fs	-.011	.005	63.792	-2.161	.034	-.022	-.001
[diagnosis_group=1] * celf_fs	.015	.006	61.257	2.355	.022	.002	.028
[diagnosis_group=2] * celf_fs	0[b]	0	.	.	.	.	.

Note. Fixed effects estimates are shown for predictors of semantic distance (SemDis). Rows marked 0[b] indicate the comparison categories for group and timepoint variables.

a. Dependent Variable: sem_dis_average.

b. This parameter is set to zero because it is redundant.

Table 11

Bivariate Correlations Among Narrative Outcomes and Cognitive-Language Variables

Variable	1	2	3	4	5	6	7	8
1. Episodic details	—							
2. AI semantic categories	.29**	—						
3. Semantic coherence (SemDis)	.09	−.02	—					
4. Word count	.66***	.34***	.07	—				
5. Expressive language (CELF-FS)	.33***	.30**	−.03	.37***	—			
6. Full Scale IQ	.43***	.28**	.10	.57***	.69***	—		
7. Vineland Communication	.23*	.15	.16	.27*	.30**	.52***	—	
8. Chronological age	.06	.04	−.04	.09	.28**	.03	−.42***	—

Note. Values are Pearson correlations. Lower SemDis scores reflect greater narrative coherence. Ns ranged from 79–100 due to pairwise deletion. *p < .05. ** p < .01. *** p < .001.

Figure 1

Normal QQ-Plot of Residuals for total_int

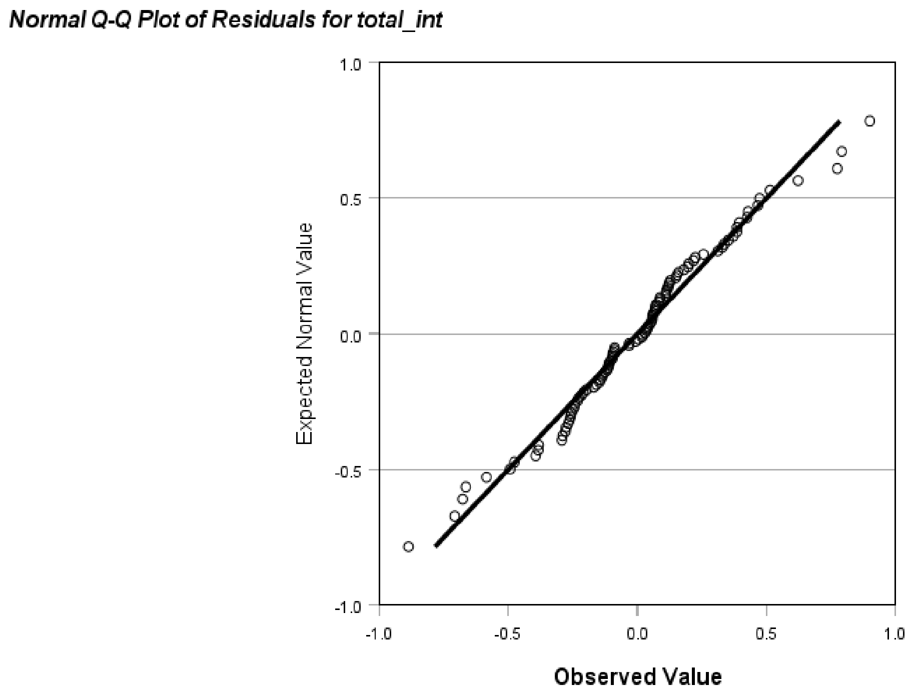

Figure 2

Detrended Normal Q-Q Plot of Residuals total_int

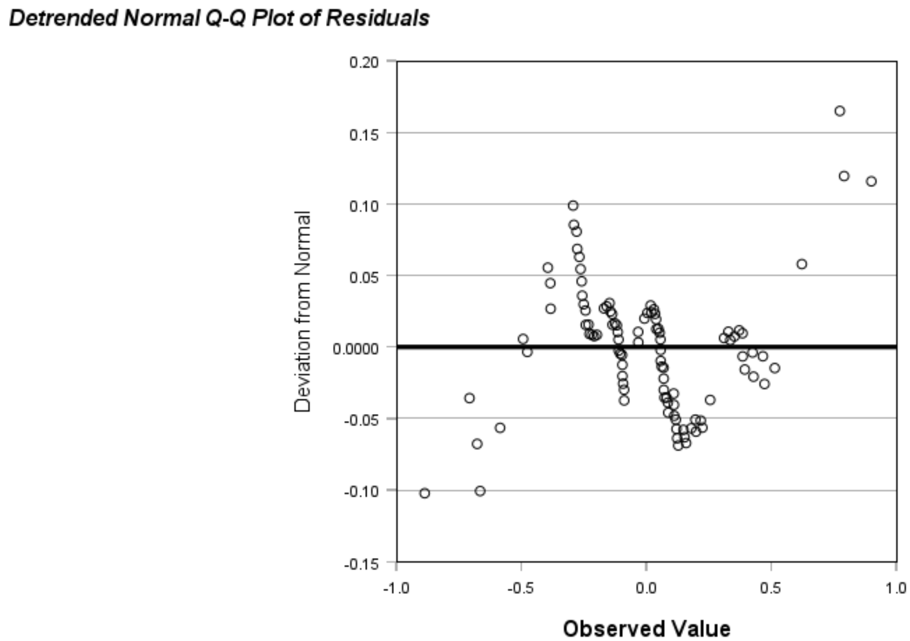

Figure 3

Predicted v Residuals

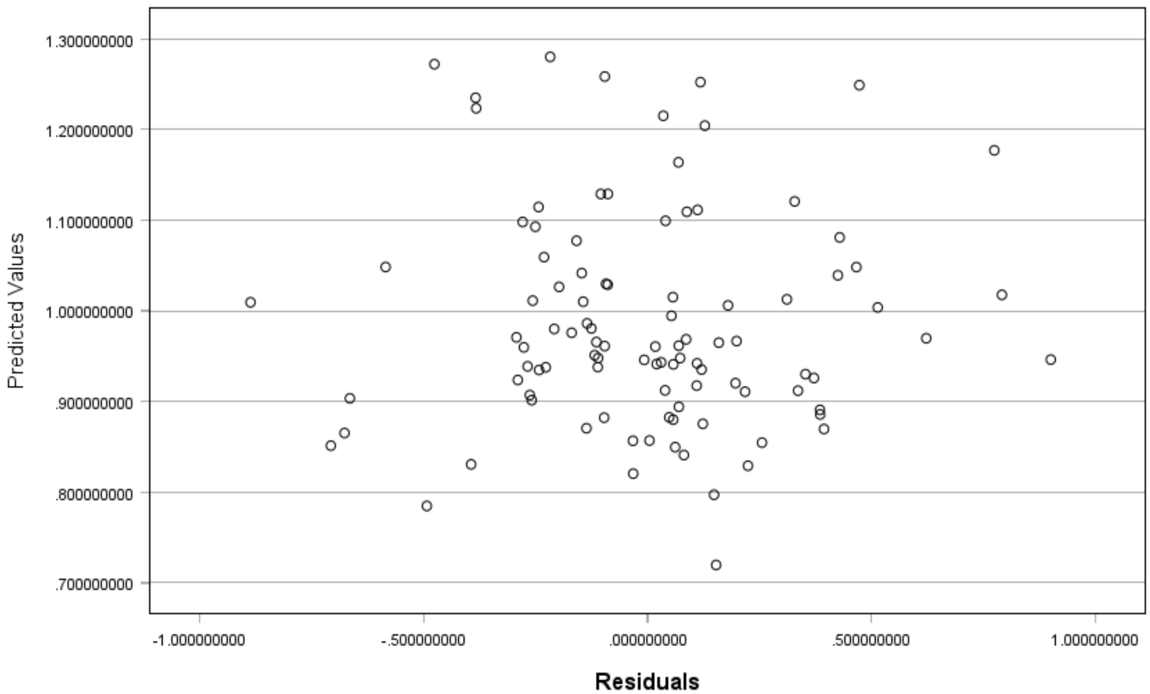

Figure 4

Normal Q-Q Plot of Residuals

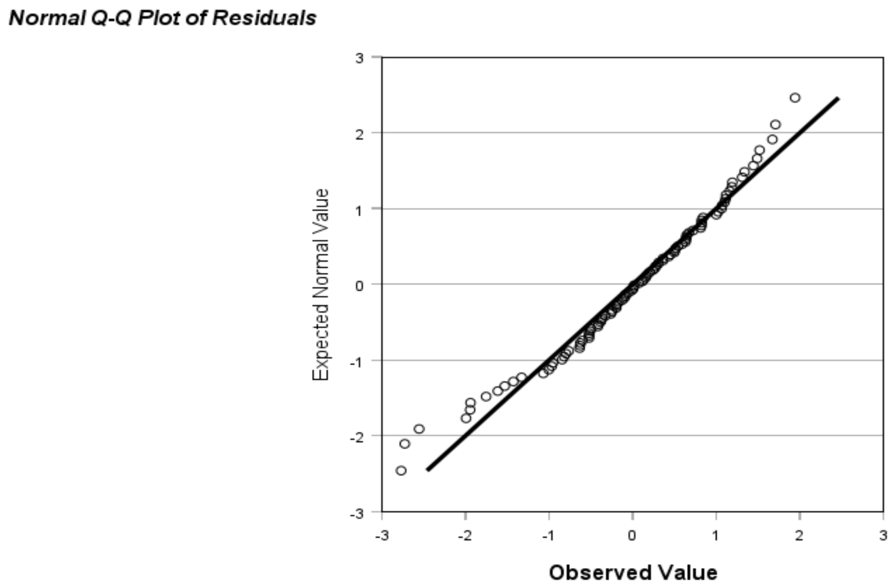

Figure 5

Detrended Normal Q-Q Plot of Residuals

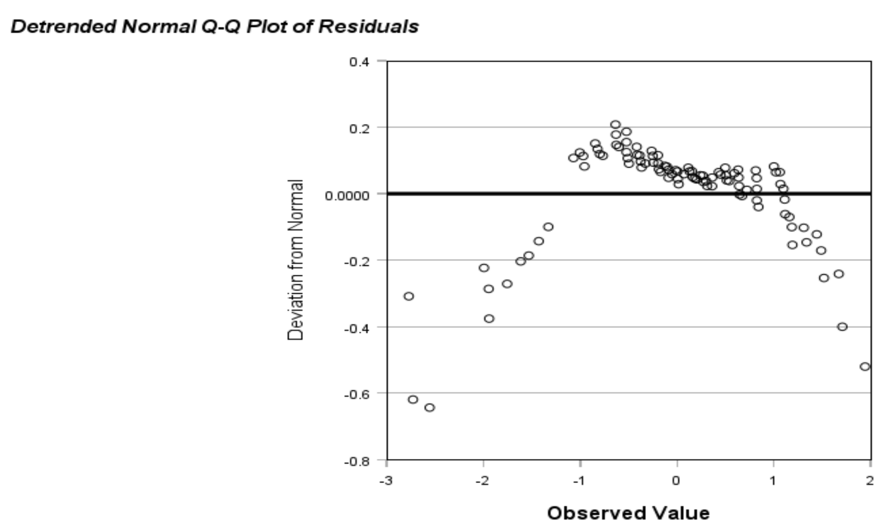

Figure 6

Predicted v Residuals

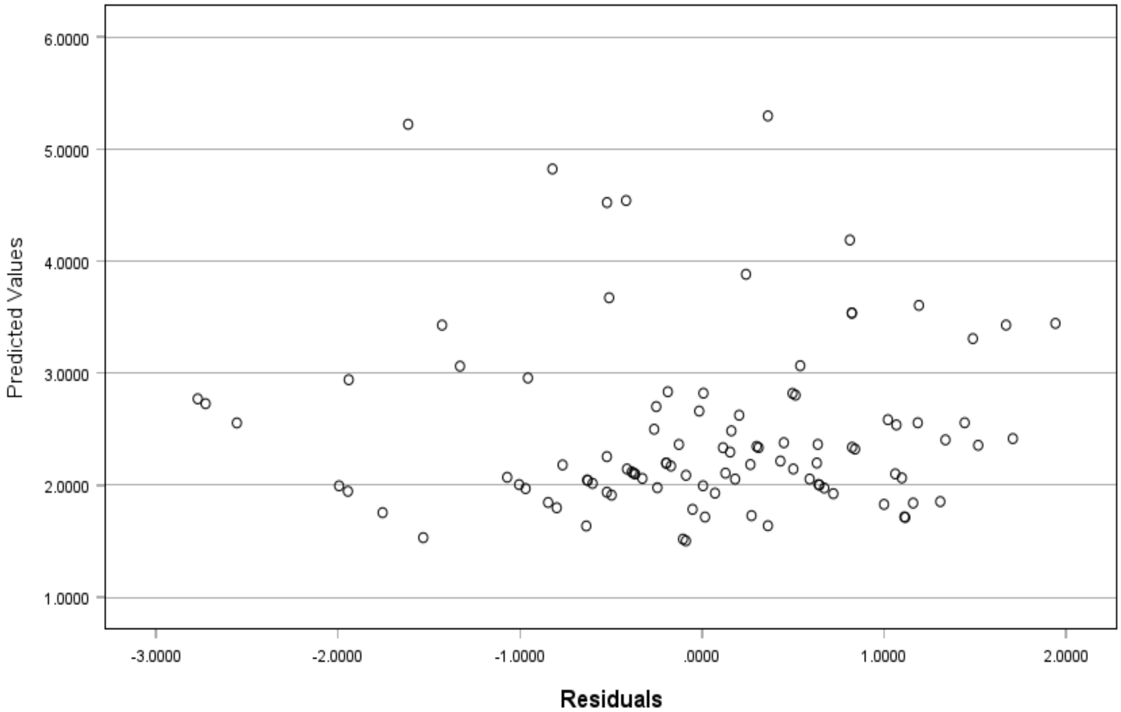

Figure 7

Normal Q-Q Plot of Residuals

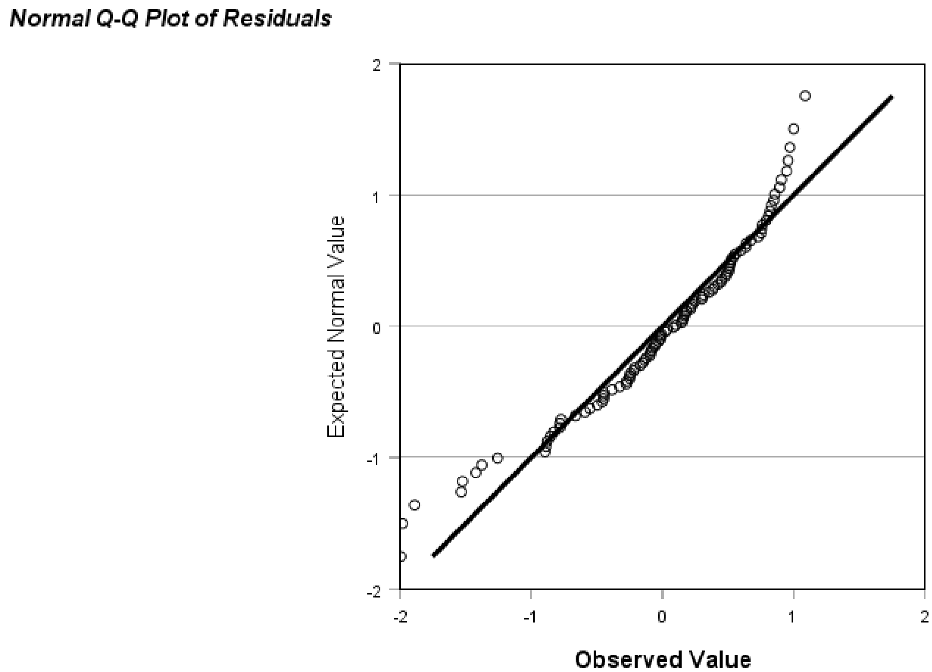

Figure 8

Detrended Normal Q-Q Plot of Residuals

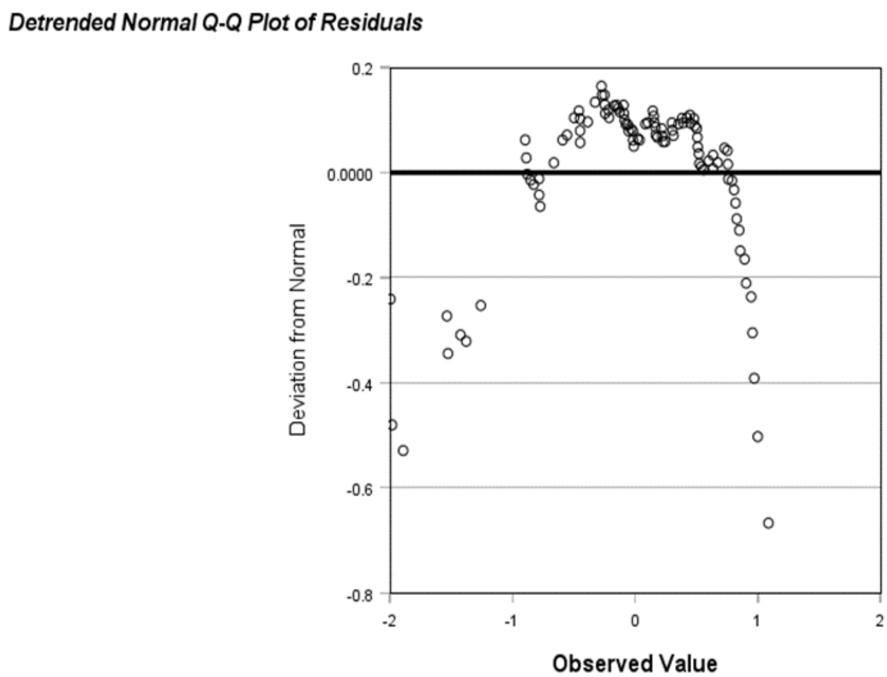

Figure 9

Predicted v Residuals

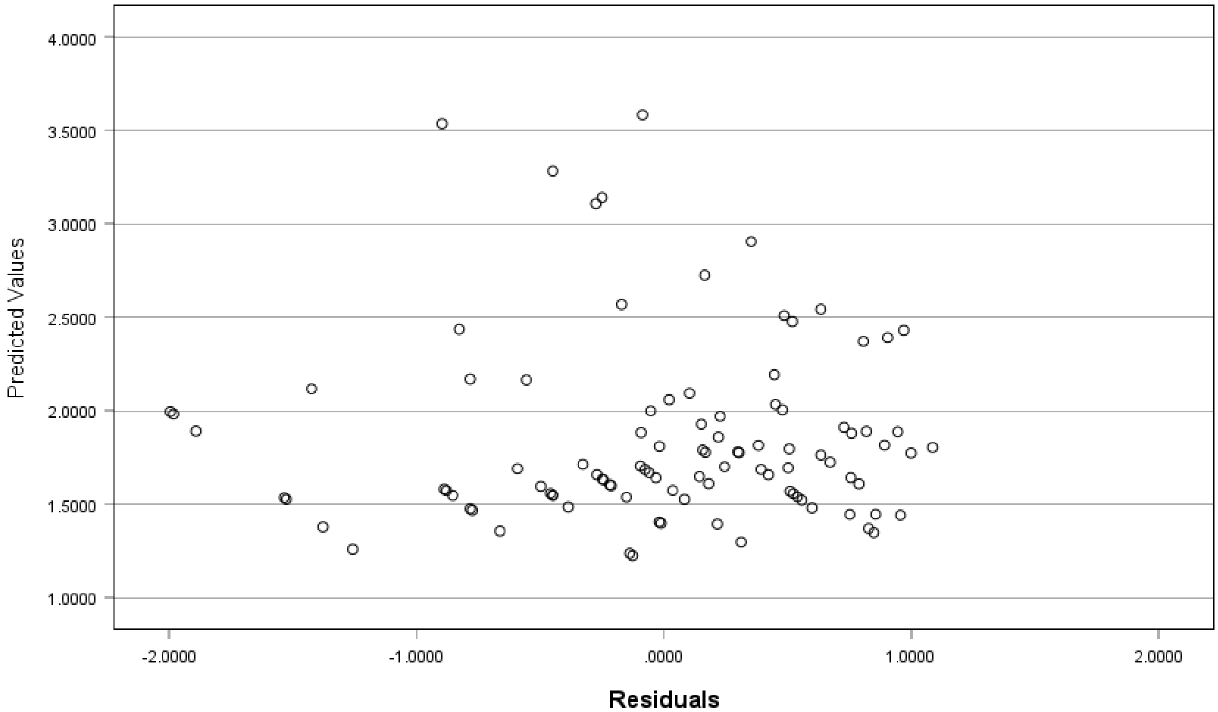

Figure 10

Histogram

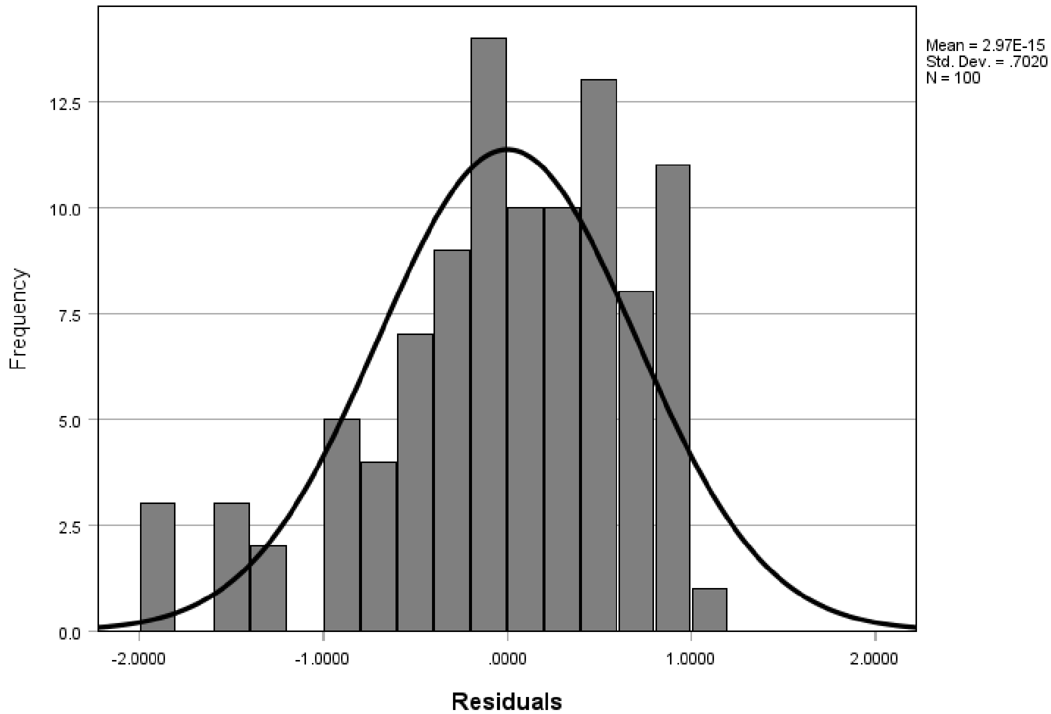

Figure 11

Normal Q-Q Plot of Residuals

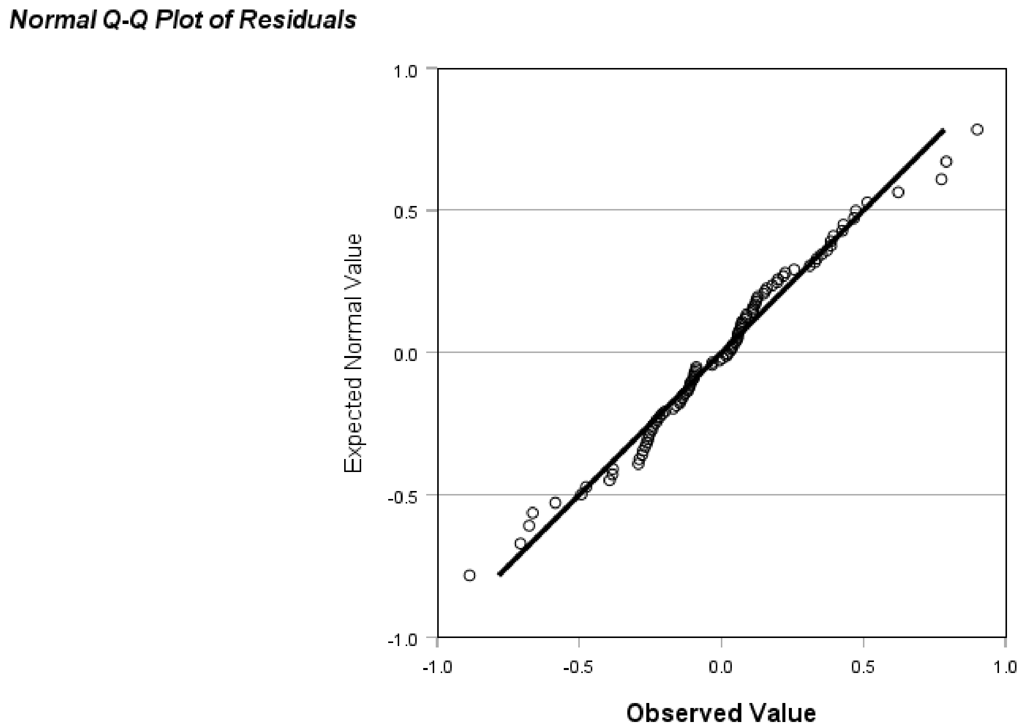

Figure 12

Detrended Normal Q-Q Plot of Residuals

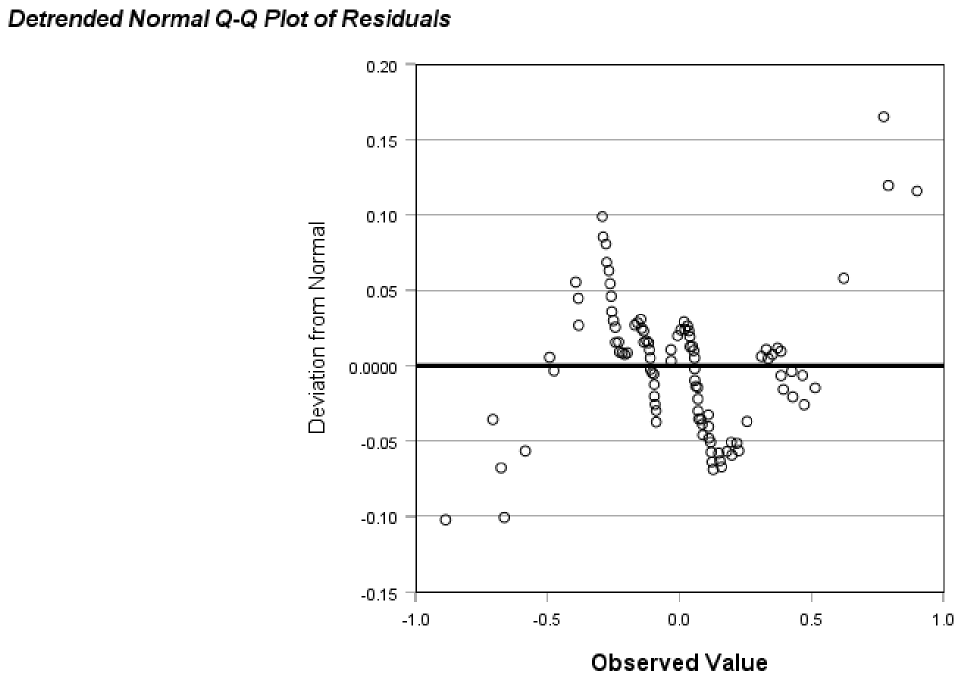

Figure 13

Predicted v Residuals Scatterplot

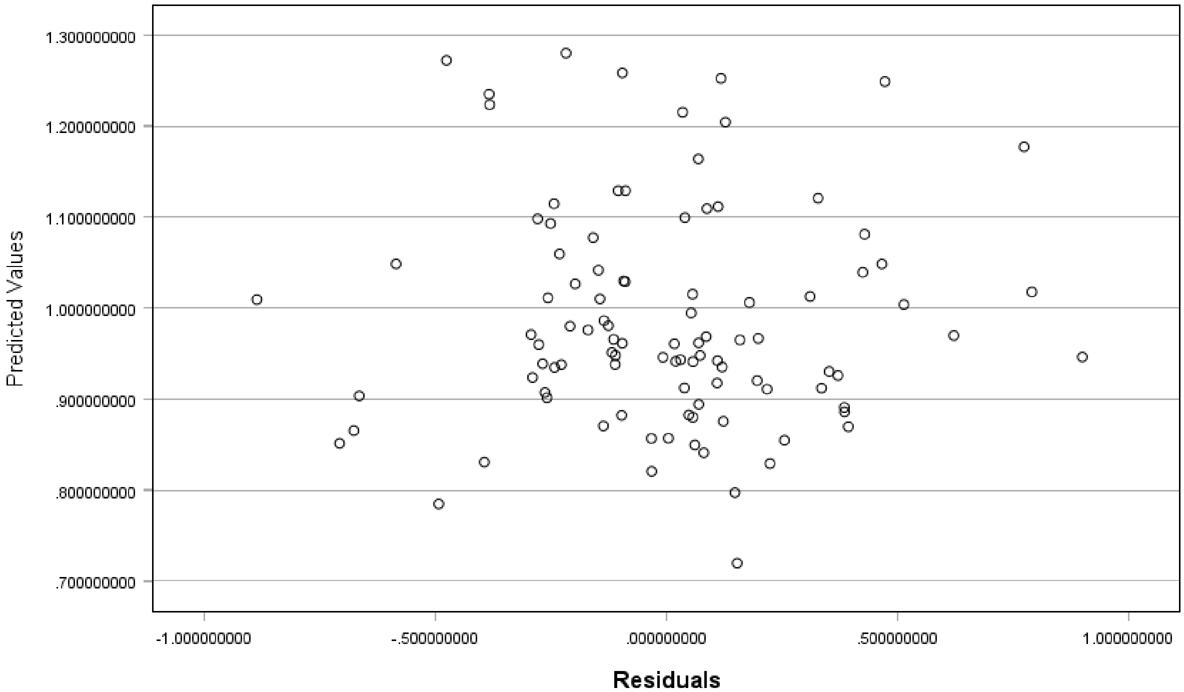

Figure 14

Histogram