Spring 2019

Assessment of Academic Vocabulary in Early Adolescents Using a Novel Sampling Method

Amber Cline
Western Kentucky University, amber.cline187@topper.wku.edu

Follow this and additional works at: https://digitalcommons.wku.edu/theses

Part of the Rehabilitation and Therapy Commons, Semantics and Pragmatics Commons, and the Syntax Commons

Recommended Citation
https://digitalcommons.wku.edu/theses/3117

This Thesis is brought to you for free and open access by TopSCHOLAR®. It has been accepted for inclusion in Masters Theses & Specialist Projects by an authorized administrator of TopSCHOLAR®. For more information, please contact topscholar@wku.edu.
ASSESSMENT OF ACADEMIC VOCABULARY IN EARLY ADOLESCENTS USING A NOVEL SAMPLING METHOD

A Thesis
Presented to
The Faculty of the Department of Communication Sciences and Disorders
Western Kentucky University
Bowling Green, KY

In Partial Fulfillment
Of the Requirements for the Degree
Master of Science

By
Amber Cline

May 2019
ASSESSMENT OF ACADEMIC VOCABULARY IN EARLY ADOLESCENTS USING A NOVEL SAMPLING METHOD

Date Recommended 3/18/2019

Janice Smith, Ph.D. Director of Thesis

Lauren E. Bland, Ph.D.

Jo Shackelford, Ed. D.

Cheryl O. Davis 4/16/19
Dean, Graduate Studies and Research Date
I dedicate this thesis to my mom. Mom, I would not have survived the thesis process, college in general, or life without you. Our daily talks and texts of encouragement and advice are what keep me going. You mean more to me than you will ever know. I love you.
ACKNOWLEDGEMENTS

First and foremost, I would like to thank my thesis committee, Drs. Janice Smith, Jo Shackelford, and Lauren Bland. All the time and dedication you all have put in to help me reach my full potential has been appreciated and valued. Dr. Smith, all the extra time and guidance you have put in to my thesis and myself can never be repaid. You have helped and guided me throughout this entire process and I would not be where I am today if not for you. You have helped me become a better researcher, a better writer, and a better speech-language pathologist.

I also want to say thank you to my wonderful parents and family for supporting me throughout this entire process. I would not have made it through without your constant support and love. Dad, thank you for all the encouraging words and conversations. Mom, thanks for everything. You are my rock and the sole reason I have persevered through my thesis and my entire education.

Thank you, Anna McKay, Taylor Nutgrass, and Shelby Jo Cecil for helping me conduct my research. You helped my data collection time go smoothly and helped me make DM-STEAM into the project we all thought it could be. A special thank you to Anna McKay for all the extra time and dedication you put into my project, all while working on yours as well. Your help did not go unnoticed and I am forever grateful.

Finally, I want to thank my friends Breanna Russell, Maggie Cartwright, and DeAnn Wallen for supporting me throughout this process. You helped me through graduate school and I know I would not have been able to complete my thesis without your support. Thank you.
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Literature Review</td>
<td>3</td>
</tr>
<tr>
<td>Methodology</td>
<td>15</td>
</tr>
<tr>
<td>Results</td>
<td>21</td>
</tr>
<tr>
<td>Discussion</td>
<td>27</td>
</tr>
<tr>
<td>Conclusion</td>
<td>31</td>
</tr>
<tr>
<td>References</td>
<td>32</td>
</tr>
<tr>
<td>Appendix A</td>
<td>38</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1. DM-STEAM Protocol ................................................................. 19

Figure 2. Path Analysis ................................................................. 26
LIST OF TABLES

Table 1. Results of Levels of Lexical Sophistication…………………………………….23
Table 2. Paired t-test Results…………………………………………………………………25
Table 3. Path Analysis Results…………………………………………………………………26
The current study examined a method of language sampling (the Dixit Method-Science, Math, Engineering, Arts, and Math) in early adolescents with typically developing language. The purpose of this study was to examine the effectiveness of the DM-STEAM in eliciting lexically sophisticated spoken language samples in the early adolescent population. To examine lexical sophistication, traditional measures of analysis such as mean length of utterance (MLU) and average type token ratio (AVG TTR) were applied along with a measure of low frequency vocabulary. To compare performance on the DM-STEAM, school standardized assessments were obtained to measure student skill in academic content areas. Twenty-two student participants in the sixth grade (11 years to 12 years 11 months) were recruited from a local elementary school. The data was evaluated using a paired tailed t test and a path analysis test. Although the sample size is small, results from the study indicate the DM-STEAM elicits low frequency academic vocabulary in early adolescent populations.
Introduction

Statement of the Problem

Speech-language pathologists working in public-school systems are tasked with helping children with language disorders be successful in the general education classroom. One aspect of success in the classroom is access to academic content and the ability to comprehend the text taught. As grade level increases, access to the content becomes more limited for children with language delays or disorders. One reason for this decrease in access is increased complexity language levels.

A particularly challenging population for speech-language pathologists (SLPs) in the schools are early adolescents. During adolescence, the requirements of language expression become increasingly complex. In this phase of language development, SLPs may opt to gear evaluations towards levels of lexical sophistication rather than just language acquisition. In general, lexical sophistication is considered the amount of advanced or specialized words students use (Kim, Crossley, & Kyle, 2017). In order to be academically successful, students are required to incorporate specialized vocabulary from specific content areas into both written and oral products. For example, a student may be asked to both write a report on the outcomes of a scientific experiment as well as orally report the findings to her class. However, evaluation of lexical sophistication for both spoken and written language in relationship to academic content areas such as science, technology, engineering, art, and math (STEAM) is limited. Specialized vocabulary for academic content areas is referred to as “academic language.”

Evaluation procedures need to adapt to ensure that language abilities during this crucial cognitive developmental stage are properly assessed. The protocols used now are
narrow in their ability to elicit and analyze spoken language samples as a whole. These limitations have resulted in a gap in the proper identification and treatment of adolescents with language disorders (Spencer, Clegg, & Stackhouse, 2010).

Students who have an unidentified disorder may not have equal access to academic content or be able to express themselves adequately, putting them at risk for academic and social delay, and in some cases, incarceration (Bryan, Freer, & Furlong, 2007). With increased focus on STEM (science, technology, engineering, and math) literacy in public schools, these content specific concepts and vocabulary must be incorporated into special education instruction to ensure students with disabilities are afforded equal access to academic content. According to The National Science Foundation (2007), students with disabilities consistently underperform in STEM coursework, emphasizing the need for valid methods of evaluating academic language (p. 1).

**Purpose of the Study**

In order to develop valid methods of evaluating academic language for those who have language disorders, these novel techniques must be tested with typically developing peers. Smith and Smith (2018) collected preliminary data on a new technique, the Dixit Method (DM), to elicit more lexically diverse language than compared to standard interview procedures. Although there were no statistically significant differences in traditional measurements (such as type-token ratio and mean length of utterance) in lexical diversity, the DM was determined to provide valuable information regarding lexical sophistication. Results, based on an ecological analysis approach, indicated that the DM elicited more representative language samples from early adolescents than a
standard interview protocol. Based on these findings, this research proposes to investigate the potential for an adaption of the DM to strategically elicit lower frequency academic vocabulary, referred to as the DM-STEAM (Dixit Method- Science, Technology, Engineering, Art, and Math).

Research Questions

The project’s primary objective will be to determine if the elicitation/collection methods, the DM and the DM-STEAM, are effective in eliciting lexically sophisticated samples reflective of adolescents’ use of academic language. As such, the following question was proposed: Do early adolescents increase lexical sophistication when specifically prompted to use academic language (DM-STEAM) in comparison to unprompted academic language (DM) elicitation?

Secondly, data was collected to determine if a relationship exists between standardized scores on K-PREP and STAR testing; language samples were also collected. For this objective, the question becomes: Is there a correlation between school assessment scores and level of lexical sophistication in early adolescents’ language samples when specifically prompted to use academic language with the DM-STEAM?

Lastly, it is hoped the results of this study will yield evidence to provide SLPs and educators with practical knowledge regarding best practices for elicitation of academic language samples from early adolescents.

LITERATURE REVIEW

Language and Lexical Sophistication

In adolescence, language becomes more sophisticated. In the field of speech language pathology, overall level of vocabulary use can be measured through lexical
sophistication and lexical diversity (Kim, Crossley, & Kyle, 2017). Researchers can agree that sophisticated words are commonly used less frequently in everyday conversation. According to Coxhead (2000), lexical sophistication and low frequency vocabulary can also be described as words that are widely used in academic contexts. According to Kim and colleagues, lexical sophistication is defined as “the proportion of relatively advanced words in the learner’s samples… [and are] generally conceptualized as low-frequency words,” (Kim, et al., 2017, p. 121.) For this study, lexical sophistication is operationalized as the number of specialized words. Specialized words are considered vocabulary related to a particular content concentration or areas of academic study.

Lexical sophistication is used in the field of speech-language pathology as a measure for academic and overall lexical success. Studies show high lexical sophistication is predictive of writing performance, writing quality, and lexical proficiency (Leki, Cumming, & Silva, 2008; Crossley et al., 2009, 2010, 2011a, 2011b, 2013; Crossley, Kyle, & McNamara, 2015; Kyle & Crossley, 2015). Multiple studies have found that proficient writers are more likely to use “low frequency words, words that occur in fewer contexts, imaginable words… and words that are acquired at a later age” (Laufer & Nation, 1995; Crossley & McNamara, 2012; Crossley et al., 2014; Jung, Crossley, & McNamara, 2015; Kyle & Crossley, 2016).

**Adolescent Language.** As adolescents grow and learn, language acquisition becomes more subtle compared to acquisition in previous stages. Language growth in adolescents is gradual and involves more sophisticated aspects of language such as fictional vocabulary like “muggle” and social mores. Language growth is significant as well as substantial during adolescence in the areas of pragmatics, syntax, and semantics.
In pragmatics, conversational skills show the most dramatic change, according to Nippold (2000). Three influential areas of change are topic selection, peer communication, and family communication (Nippold, 2000). Topic selections become more personal for adolescents. Topics tend to be driven by personal experiences and concerns. Adolescents are more likely to discuss intensely personal issues, such as embarrassing experiences or divorce, with peers and family members.

Peer communication is measured by the amount of time engaged in social interaction with peers. Studies have shown that as age and grade level increase, the amount of peer interaction increases significantly (Raffaelli & Duckett, 1989). Peer interactions are crucial for the happiness and personal wellbeing of adolescents. Interactions ease the transitional experiences such as puberty, increased academic responsibilities, and more complex family dynamics of adolescents (Raffaelli & Duckett, 1989). Family communication is another important element as these relationships provide support that cannot be given by peers.

Syntax is another area of growth for adolescents. During this period, the use of formal written language is required for academic success. For example, adolescents are required to use more complex sentence structure in order to write academic papers on a variety of topics for schools. Adolescent syntax growth has been measured using persuasive writing samples (Crowhurst, 1980a, 1980b, 1987; Crowhurst & Piche, 1979, Rubin & Piche, 1979). Persuasive writing is considered cognitively and linguistically more demanding compared to other types of writing and requires higher level thinking skills (Crowhurst & Piche, 1979). With higher level writing, there is a need for longer sentences containing more clausal subordination and use of dependent clauses. Better
links between sentences also develop as evidenced by increased use of connectives such as conjunctions and coordinating clauses. Adverbial connectors, (i.e. otherwise, therefore, however), may also be used to accomplish higher levels of sophistication (Crowhurst, 1980a, 1980b, 1987; Crowhurst & Piche, 1979, Rubin & Piche, 1979). Development of syntax skill results in higher level writing proficiency.

Finally, semantic skills also change significantly. During adolescence, semantic skill growth is observed in the understanding and use of figurative language. Figurative language includes metaphors, idioms, similes, and proverbs. Proverbs are considered to be the most sophisticated type of figurative language due to their increased abstract content. Additionally, proverbs are the last form of figurative language to be mastered. Proverbs can be difficult to understand because of abstract contexts such as hidden meaning or “moral of the story” used to connect a fictionalized story to a deeper present day meaning. Relatability and expository type, like poetic style, inference and voice, may also “challenge the interpretative abilities of individuals” (Nippold, 2000, p.26). Evidence has shown proverbs are easier to interpret when they include familiar vocabulary and depict moral standards (i.e. murder is bad) rather than attitudes and judgements (i.e. cheating results in feelings of shame) and included familiar vocabulary, are easier to interpret. Lastly, Nippold purports proverbs become easier to understand as adolescents obtain supportive linguistic contexts to facilitate comprehension.

**Academic Language**

Public school systems are acknowledging the need to instruct students on academic language. The Common Core State Standards now reflect the need to teach academic language by using language objectives alongside existing content objectives.
The language objectives assist teachers in identifying language needs of lessons taught, academic language presented, and strategies for instruction (Lindahl & Watkins, 2014). According to Uccelli and Galloway (2017), learning academic language is crucial for success in school and beyond. They stated that “academic language proficiency also supports citizens’ access to public information, such as health advice or political news, and facilitates civic participation” (p. 399). Academic language improves academic success and establishes a foundation to become a productive member of society. A standard definition of “academic language” has not been determined. According to Friedberg, Mitchell, and Brook (2017), academic language is defined as “formal English rules, structure, and content for academic dialogue and text, and the communicative conventions that allow students to meet the demands of school environments” (p. 2).

Lindahl and Watkins identified aspects of academic language: vocabulary, functional language, grammar, word study, reading comprehension, and writing/conventions. Vocabulary demands include “acquiring new words that represent known concepts, acquiring new words that represent new concepts, clarifying and enriching the meaning of new words…or using context clues to decipher meaning” (Lindahl & Watkins, 2014, P. 198). These demands encompass the specific vocabulary of content areas (i.e. science, math) and the ability to use them correctly. Students are required to be proficient when using academic vocabulary orally and in written form. Grammar and word study are also important for academic language proficiency. Students integrate both to appropriately express the structure of language and words. Knowledge of grammar provides information regarding students’ word awareness and meaning (Lindahl & Watkins, 2014).
**Academic Language Versus Social Language.** Social language is used every day to express basic wants and needs. This aspect of language is conversational in nature. Academic language differs greatly from social language in that it is “decontextualized, more complex, more abstract, and places higher demands on student cognition” (Lindahl & Watkins, 2014, p.199). Social language is considered to consist of high frequency words whereas academic language often utilizes lower frequency words. Academic language is denser in terms of meaningful vocabulary and pertinence of information. Learning academic language is influenced by “cultural and experiential demands” or background and experiences of the learners (Lindahl & Watkins, 2014). Social language may be limited by personal surroundings and language learned in the home. In contrast, academic language incorporates content terminology from a multitude of backgrounds for a diverse learning experience.

**Academic Language Assessments.** There are students who are below age level in academic language performance. This shows a need for evaluations to help support those students. Currently, there is little research in the area of academic language assessment. This type of assessment is needed in order to identify students who need additional support to learn academic language. Uccelli and Galloway (2017) completed research in the area of academic language assessments. The Core Academic Language Skills (CALS) was created by Uccelli and Galloway and identifies seven domains that need to be assessed to measure a student’s academic language level and proficiency. The domains include: unpacking dense information, connecting ideas logically, tracking participants and themes, organizing analytic texts, understanding metalinguistic vocabulary, interpreting writers’ viewpoints, and recognizing academic texts. These domains were
used to create the *Core Academic Language Skills Instrument* (CALS-I). The CALS-I can be administered in 50 minutes and uses receptive and expressive multiple-choice questions, matching items, and short written responses (Uccelli & Galloway, 2017). Subjects included 218 students in grades 4-6 primarily coming from lower socioeconomic status SES households.

This study using the CALS-I showed significant differences in scores based on grade level and SES. Scores were higher with more advanced grade levels, showing an increase in academic language as students progressed in grade level. Scores were significantly lower for students of lower SES and students who qualified for free and reduced lunch. These findings highlight the need for equal access to academic language instruction experiences, regardless of grade level or SES. A positive correlation was found between CALS-I scores and reading comprehension; as the CALS-I scores increased so did reading comprehension scores. This implies academic language knowledge is essential for success when reading academic texts.

Uccelli and Galloway (2017) also administered three other assessments with subjects to measure their academic language. The *Gates-MacGinitie Reading Test* was used to test reading comprehension; the *Test of Silent Word Reading Fluency-2* to test word reading fluency; and the *Vocabulary Association Test* was used to measure academic vocabulary depth. All of the assessments used in the studies measured an academic language area and provided normative data to academic language instruction. Receptive measures are important, but there is a need for assessments of expressive academic language skills as these are the skills required in the classroom.

**Language Sampling**
Language samples are an important part of assessing a student’s overall language ability. A language sample can provide information about a student’s vocabulary, mean length of utterance, type token ratio, and spoken language ability. Language sampling is used to measure a student’s strengths and weaknesses by scoring the complexity of the sample and is usually 50-100 utterances in length to elicit enough information to be scored.

Language samples are informal measures that can be completed in a variety of ways. The most common elicitation techniques are freeplay, interviews, and narratives. For adolescents, the two most commonly used methods are interviews and narratives. Interviews and narratives tend to elicit more complex language samples that are reflective of age, grade level, and lexical sophistication.

**Interviews.** An interview, also referred to as reciprocal conversation, is used as an informal language sampling method. An interview usually consists of a turn taking task involving a question or conversation starter, such as “Tell me about your family” (Nippold, 2000). For a successful interview, the evidence suggests a child directed approach is most effective, wherein the interviewer introduces the topic and the child steers the conversation (Southwood & Russell, 2004).

According to Evans and Craig (1992), interviews produce more utterances, longer utterances, more semantically and syntactically complex forms, more simultaneous speech, and higher levels of responsiveness. They also outline significant differences between interviews and other language sampling techniques stating less time is required without “(a) omitting behaviors observed; (b) jeopardizing conversational validity the way more higher structured tasks do; and (c) affecting the diagnostic validity of structural
assessment tools” (Evans & Craig, 1992, p. 351). According to Mirsaleh, Abdi, Rezai, and Kashani (2011), interviews also produce higher type token ratios compared to other sampling techniques. Most interview protocols are 15-30 minutes in length and involve a topic of interest. Reduced structure and students’ decreased feeling of stress during interviews are an advantage.

**Narratives.** Narrative methods of language sampling are commonly used with adolescent populations because integration of language knowledge and application are required to produce a narrative allowing for the evaluation of language complexity. Narratives require higher syntactic and sematic complexity and typically produce longer samples than other techniques, including interviews (Nippold et al., 2014). Two subtypes of narratives are discussed in the literature: story generation and story retell.

Both story generation and story retell are defined as “a monologue used to convey information and may involve providing factual descriptions of explanations of events” (Westerveld & Moran, 2013, p. 727). Story generation refers to the telling of a story using one’s own information and facts and is often elicited using questions or prompts. Story retell differs in that it involves listening to a story then restating it including vital information and plot.

Evidence has shown narrative techniques have advantages compared to other forms of elicitation. According to Mirsaleh, Abdi, Rezai, and Kashani (2011), narratives produce higher mean length of utterance compared to other techniques such as freeplay. When comparing story retell to story generation, story retell produces longer utterances whereas story generation produces longer samples (Nippold et al., 2014). Even though there is a difference, both methods still elicit more complex utterances and samples in
comparison to interviews. According to Nippold and colleagues (2014), narrative tasks elicited utterances were “5.43 words longer and 81% denser” when compared to interviews (p. 880).

**Dixit Method.** During the past year, research was conducted in the area of language sampling. Smith and Smith (2018) collected preliminary data on a novel technique, the Dixit Method (DM), to elicit more lexically diverse language than standard interview procedures. The DM utilizes richly illustrated cards from the game Dixit to elicit language samples (board game: Roubira & Cardouat, 2008). The researchers also applied an ecological approach, proposed by linguist Jarvis in 2013, to analyze lexical diversity.

Two language elicitation methods were compared: interview and the DM. Their analysis was completed and assessed for standard measures, type token ratio (TTR), average type token ratio (AVG TTR), and mean length of utterance (MLU). Their ecological analysis approach evaluated six of the seven aspects proposed by Jarvis (2013): size, richness, evenness, disparity, importance, and dispersion. This ecological approach was hypothesized to provide a more structured and detailed analysis of the results from adolescent language samples.

Results supported the assertion that the DM produces more lexically diverse language sampling approach than the interview method. Further, though evidence was very preliminary, the ecological analysis approach appeared to better reflect representative features of lexical complexity (low frequency vocabulary, specialized language, abstract concepts) for early adolescents. However, based on traditional measures of language diversity (TTR, AVG TTR, and MLU), the results were unclear revealing no statistically significant difference between the two techniques. According to
Smith (2018), “the current traditional methods… are not comprehensive enough to authentically reflect spoken language performance in early adolescents” (p. 34). For example, results from the ecological analysis, reveal significantly higher size and richness for the DM than the standard interview protocol. Smith (2018) states, “…the samples were significantly longer and students utilized unique words and word types more often” (p. 36). In addition, students used vocabulary well beyond what was elicited during the interview. A third aspect, importance, is defined as the frequency with which words were used. Results showed the DM elicited more specialized and unique vocabulary. Measurement of evenness and disparity, comparison of amount of word types used and relationship between word types, were found to reflect lexical diversity expected of early adolescent children. No significant difference was found between methods for dispersion rate of words (Smith, 2018). These findings support using the DM to elicit more representative language samples for early adolescents compared to the standard interview. Lastly, the results support applying an ecological approach to analyzing lexical diversity provides a more holistic perspective than traditional measures.

**Current Study**

The purpose of this research was to investigate the effectiveness of an adapted version of the DM to strategically elicit lower frequency academic vocabulary. For this adapted procedure, illustrated cards were preselected to improve the likelihood of eliciting vocabulary respective to STEAM content concentrations. The Dixit Method-STEAM adapted version (DM-STEAM) implements targeted elicitation procedures to capture low frequency vocabulary. Analysis of results will examine potential impacts of
elicitation procedures on the level of lexical sophistication and academic vocabulary use in adolescents.

The primary objective was to determine if the DM and the DM-STEAM were effective in eliciting lexically sophisticated samples that reflect early adolescents’ use of academic language. Lexical sophistication, used to measure language ability for specific age groups, may help speech-language pathologists to identify and treat adolescents with language disorders. Level of lexical sophistication was measured to reflect use of academic language and to indicate overall academic success.

The second objective was to determine if a relationship exists between standardized scores (i.e. STAR reading and math assessments and K-PREP state-wide assessments) and collected language samples. The third objective was to advance the knowledge in the field regarding best practices for elicitation and collection of academic language samples for early adolescents. To fulfill this purpose, the study sought to answer the following research questions:

1. Do early adolescents show increased lexical sophistication when specifically prompted to use academic language (DM-STEAM) in comparison to unprompted academic language (DM) elicitation?
2. Is there a correlation between school standardized assessment scores and level of lexical sophistication in early adolescents’ language samples when specifically prompted to use academic language with the DM-STEAM?

It was hypothesized that adolescents would respond with increased lexically sophisticated language during administration of the DM-STEAM. Sophisticated responses are expected
due to expected levels of cognitive development and academic vocabulary knowledge associated with early adolescent development.

**Methodology**

**Participants**

Thirty-three students participated in this study. Participants included 20 males and 13 females and were between the ages of 11 years and 12 years 11 months. Twenty-two samples were transcribed for analysis. All of the participants were recruited from a local elementary school and all lived in the same community in south central Kentucky.

The study was approved by the Western Kentucky University Institutional Review Board prior to recruitment and data collection. A parent or guardian had to sign a written consent form giving permission for student participation. Data collection was conducted during pre-arranged, teacher preferred times. Participants were given small incentives for participating: a snack, drink, and fidget cube.

During the data collection session, participants met with a research assistant in a one-on-one setting. Research assistants were trained by the primary investigator in the language sampling protocol and administration. Training included observing an administration and conducting a session prior to the beginning of data collection. All data collectors were from the same accredited university: one was an American Speech-Language-Hearing Association certified speech-language pathologist who is a faculty member at the university, one was a graduate student in the speech-language pathology program, and the other two data collectors were undergraduate students in the communication sciences and disorders program. All data collectors were CITI (Collaborative Institutional Training Initiative) trained prior to initiation of the study.
Materials

Prior to the start of data collection, the researchers distributed printed explanations of the study and consent forms to teachers at the school. Teachers then sent the consent forms home with participants to be signed by a parent or guardian. After language samples were collected, participants’ homeroom teachers were asked to fill out the Student Language Scales (SLS) from the *Test of Integrated Language and Literacy Skills (TILLS)* (Nelson, Howes, & Anderson, 2016). For each participant, the homeroom teacher served as the primary teacher and completed the SLS screener. The SLS is a short language screener used to compare a student’s academic success to their same age peers (Nelson, Howes, & Anderson, 2016). Primary teachers answered 12 questions regarding language performance using a 1-8 rating scale with 1 being the lowest score possible. Students who were scored a 2 or lower in two or more areas on the scale were considered at significant risk of language delay or disorders and were referred to their primary teacher for further evaluation.

Data collectors were trained to use the novel language sampling technique to elicit academic vocabulary, the Dixit Method-Science, Technology, Engineering, Arts, and Math (DM-STEAM), which was adapted from the Dixit Method (Smith & Smith, 2018). The protocol outlined procedures to specifically elicit academic vocabulary using illustrated cards from the Dixit game (refer to DM-STEAM protocol in *Figure 1*).

Each data collector used a DM-STEAM deck of preselected content specific Dixit cards. Protocol form were used to standardize data collection during academic language sampling. Each deck had 25 preselected cards that reflected the five content areas with five cards representing each content area. Cards were chosen from two Dixit game decks,
The Dixit Original Game and Dixit Memories Expansion card pack. The Dixit game publisher, Libellud, gave permission for the use of the cards in this study. For examples of the illustrated Dixit cards, please refer to Appendix A.

GoPro cameras and stands were utilized to record each session. Each data collector also had color coded note pads, which were used for the brainstorming portion of the protocol. The note pads were color coded according to the content area (blue-science, green-technology, pink-engineering, orange-the arts, neon pink-math).

**Procedures**

Data collection took place at the elementary school where participants were recruited. Teachers were invited to observe sessions and one teacher chose to do so. Sessions were conducted in a classroom with tables and chairs arranged into testing stations. On one data collection day, the primary location was not available, so an alternate location was provided. This location was a stage situated between a gymnasium and lunchroom. Though there was an increased noise level, tables and chairs were arranged into similar testing stations. Each participant was scheduled for a 45-minute session to ensure enough time for completion of the protocol. Prior to completing data collection, researchers contacted primary teachers to ensure availability of students.

Consent documentation was signed by parents or guardians of participants prior to data collection sessions. In addition to consent document, data collectors also required an assent document to be signed by participants. At the beginning of the session, data collectors read aloud the assent form to participants informing them of their rights for participation in the study and to confirm their willingness to participate and be videoed.
Participants then signed and initialed the document. After assent forms was signed, data collectors conducted the DM-STEAM protocol to elicit language samples.

The DM-STEAM protocol consisted of a brainstorming round, a sorting round, and then five rounds of describing and telling a short story by integrating two or three illustrated cards. During the protocol, sessions were primarily student-driven based on preferences of cards and academic content areas. Students were given minimal prompting throughout data collection. Data collectors did not communicate preference or perceived accuracy during card sorting. For example, students were instructed, “Cards can go in a lot of places; we want to see where 6th graders will put them” to avoid priming participants with content specific language. Data collectors did not provide vocabulary examples unless prompted by the student either verbally or through behaviors such as eye contact, vocal expressions of frustration, or other nonverbal cues. When prompting was needed, all data collectors used the same generic prompt and vocabulary target. For example, a child struggling to generate vocabulary regarding engineering was prompted with the statement, “Building things is a part of engineering.” After data collection sessions were complete, students were given their incentives and returned to class.

At the conclusion of data collection for the DM-STEAM protocol, collected language samples collected were reviewed and transcribed and verified by the primary investigator, faculty mentor, and the research team. Transcriptions were entered into the Systematic Analysis of Language Transcripts (SALT) 16 Research Software (Miller & Iglesias, 2012) and the Tool for the Automatic Analysis of Lexical Sophistication (TAALES) (Kyle, Crossley & Berger, 2017).
**Set Up:**
Divide table into 5 categories and label STEAM (science, technology, engineering, (the) arts, and math.

**Step One:**
- “I want you to brainstorm 5 words about each category and write it on the label provided”
- If the student gets stuck… “these words can be words used, types of (subject), and things used in (subject)”
- Prompts: say these exact so all of us use the same prompts!
  - Science: a type of science is biology
  - Technology: a word could be internet
  - Engineering: building things is a part of engineering
  - Arts: the word “songs” could be in arts
  - Math: “equations” is a word used in math

**Step Two:**
- “Now I want you to put each card into the category you think is the best fit. The cards can go in many different categories, but there needs to be at least 3 cards in each category”
- flip cards and place them flat on the table
- if too log of a pause/ thinking too hard… “remember, we are interested in your opinion”

**Step Three:**
- “If you are second guessing any, you can rearrange if you would like and if you do not have 3 in a category please rearrange to make all of them at least 3”
- pick starting deck
  - “so, which of these are your favorite?” (if more than 3 cards in favorite category, use it)
  - “do you want to pick first or me?”

**Step Four:**
- “Now I want you to describe each card using words related to (subject). You can use anything you see, anything you remember, and anything that connects to (subject)”
- researcher says “please try to say at least 3-5 (subject) words about each card”
- if student is struggling… “if you are having trouble, look back at the words you brainstormed”

**Step Five:**
- lay out cards- 2 rows
- “Now, I want you to tell me a little more about these cards. But before you do, I want you to pick 3 cards that you want to talk about or those cards you don’t want to talk about and we will get rid of those”

**Step Six:**
- using 3 cards
  - “now I want you to make a short story using the words you brainstormed or any new words you can think of related to (subject)"
  - “I want you to connect at least 2 of the cards but you can use all 3 if you want”
  - “when you are ready…”
  - if stuck… “it’s okay, just do your best”

**Step Seven:**
- sit category just done to the side!
- determine which deck will be used for the next round
  - if student seems decisive, let them pick the next deck
  - if student seems indecisive, you pick the rest of the categories
- repeat steps 4-6

---

*Figure 1. The DM-STEAM Protocol.*

**Measures**
Languages samples collected from the DM-STEAM protocol were transcribed and analyzed to determine amount of low-frequency vocabulary used by participants. Samples were evaluated using two standard measures: mean length of utterance (MLU) and average type token ratio (AVG TTR). Two portions of the DM-STEAM were measured: 1) participant descriptions of illustrated cards when prompted to use content specific words and 2) participant generated short stories where content specific information from two to three cards was integrated. Type-token ratio and average type token ratio were calculated for participant card descriptions and generated short stories (combined). Mean length of utterance was calculated for only the generated short story portion. MLU was not measured for card description responses based on the implied length limitations associated with “brainstorming.”

In order to assess MLU, samples were transcribed and the total number of utterances was calculated as well as the number of morphemes in each utterance. AVG TTR was analyzed in order to assess the level of lexical diversity in a sample. AVG TTR was analyzed by determining the number of different words elicited in comparison to the total number of words.

All language samples were transcribed and analyzed. The Tool for the Automatic Analysis of Lexical Sophistication (TAALES) was used to measure lexical sophistication (Kyle, Crossley, & Berger, 2017). TAALES analyses language samples by determining the frequency each word is used in the 8,388 American television shows and films. When analysis is complete, words are ranked by how frequently they appear in the SUBTLEXus database. The SUBTLEXus is a database that is composed of scripts of 8,388 American television shows and films.
STAR reading scores and K-PREP scores were obtained for each participant. Scores on school standard assessments and language samples were evaluated to determine if correlations existed between the two data sets. STAR reading assesses skills such as students’ word knowledge, comprehension strategies and constructing meaning, analyzing literary text, understanding author’s craft, analyzing arguments and evaluation (Renaissance Learning, 2015). K-PREP scores were also obtained for participants from the previous fifth grade school year. Fifth graders are assessed in the areas of reading, mathematics, social studies, and on-demand writing.

Results

The results answer the study’s two research questions. First, do early adolescents show increased lexical sophistication when specifically prompted to use academic language (DM-STEAM) in comparison to unprompted academic language (DM) elicitation? Second, is there a correlation between school assessment scores and level of lexical sophistication in early adolescents’ language samples when specifically prompted to use academic language with the DM-STEAM? The results are discussed below.

Thirty-three students between the ages of 11 years and 12 years 11 months participated in this study. Participants included 20 males and 13 females. Twenty-two samples were transcribed for analysis. Seven student samples were excluded due to sound quality of the sample that prohibited accurate transcribing. Three students were excluded from data analysis due to scores on the (TILLS) Student Language Scale screening from the Test of Integrated Language and Literacy Skills that indicated possible language learning difficulties. Therefore, twenty-two samples were transcribed for analysis to
determine if the DM-STEAM was more effective in eliciting more lexically sophisticated language.

**Level of Lexical Sophistication Based on Traditional Measures**

**Average type token ratio (AVG TTR).** Twenty-two language samples were analyzed for the DM-STEAM. Each sample was age and gender matched to an original DM sample from a previous study (Smith, 2018). Original DM samples refer to the DM protocol where children were allowed to use any language they wanted without prompting whereas the DM-STEAM protocol allowed children to use any language related to academic content that was specifically prompted. AVG TTR, which accounts for the size of the sample, was determined for both methods. For the DM-STEAM, AVG TTR ranged from a minimum of 0.55 to a maximum of 0.65 with a mean of 0.5973. For the (original) unprompted DM samples, AVG TTR ranged from a minimum of 0.56 to a maximum of 0.72 with a mean of 0.6050, indicating that children use the same amount of different words, no matter what protocol was used.

**Mean length of utterance.** Mean Length of Utterance (MLU) was determined for the DM-STEAM and the DM methods. For the DM-STEAM, MLU ranged from a minimum of 9.76 to a maximum of 14.14 with a mean of 11.9232. For the DM, MLU ranged from a minimum of 4.03 to a maximum of 12.22 with a mean of 7.9568, indicating that children used longer sentences when completing the DM-STEAM protocol.

**Level of Lexical Sophistication Based on Low Frequency**

**Low Frequency Vocabulary.** Low frequency vocabulary refers to the frequency in which words in a sample occur in a language as a whole. The SUBTLEXus corpus was
used to measure low frequency vocabulary in the prompted and unprompted samples (Kyle & Crossley, 2014). There are 51 million words in the SUBLTEXus, therefore, the lower the result, the more low frequency vocabulary was used in the sample. Low frequency vocabulary for the DM-STEAM ranged from 234602.8 to 436007.9 with a mean of 372127.7565. The DM ranged from 263254.1 to 398373.1 with a mean of 340782.7435, indicating that children use more specialized vocabulary with the DM.

Table One

*Results of Levels of Lexical Sophistication*

<table>
<thead>
<tr>
<th></th>
<th>DM- STEAM</th>
<th></th>
<th>DM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>min</td>
<td>max</td>
</tr>
<tr>
<td>A V G TTR</td>
<td>22</td>
<td>0.55</td>
<td>0.65</td>
</tr>
<tr>
<td>Low Frequency</td>
<td>22</td>
<td>234602.8</td>
<td>436007.9</td>
</tr>
</tbody>
</table>

Note: n=sample size; AVG TTR=average type token ratio; MLU=mean length of utterance

**Correlation Between Standardized Assessments and the DM-STEAM**

**K-PREP.** K-PREP is a school standardized test given during some grade levels in Kentucky public schools. For this study, K-PREP scores were obtained from the previous fifth grade school year. Sections for the fifth grade K-PREP test include reading, math,
social studies, and on-demand writing. Results showed significant correlation between K-PREP Math, MLU, and AVG TTR (p=0.001). There was also a significant correlation between K-PREP Social Studies and MLU (p=0.001). A significant correlation was found between K-PREP Writing, MLU, and AVG TTR (p=0.015, 0.021). Lastly, no significant correlation was found between K-PREP Reading, MLU, and AVG TTR.

**STAR Reading and Math.** STAR Reading and Math standardized assessments are administered annually. For this study, STAR Reading and Math scores for the sixth grade school year were obtained. There was a significant correlation between STAR Reading, MLU, and AVG TTR (p=0.020, 0.001, 0.002). There was also a significant correlation between STAR Math, MLU, and AVG TTR (p=0.009, 0.001, 0.001).

**Test of Significance**

A test of significance was performed to determine if there was adequate evidence to determine a significant difference between DM-STEAM and DM as well as to determine if there was a significant correlation between school standardized assessments and DM-STEAM results. To address the first research question, a paired t-test was completed. A path analysis was utilized to address the second research question.

Results for AVG TTR indicate that the mean performance for the DM-STEAM ($M = 0.5973$) was not significant compared to DM ($M = 0.6050$) with a $t$-value of -0.70 and $p$-value of 0.4889. Results for MLU indicate that the mean performance for the DM-STEAM ($M = 11.9232$) was significant compared to DM ($M = 7.9568$) with a $t$-value of 6.12 and a $p$-value of <0.001. Results for low frequency vocabulary indicate that there is a significant difference for DM-STEAM compared to the DM with a $t$-value of 2.345 and $p$-value of 0.028.
Table 2

*Paired t-test Results*

<table>
<thead>
<tr>
<th></th>
<th>DM-STEAM Mean</th>
<th>DM Mean</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLU</td>
<td>11.9232</td>
<td>7.9568</td>
<td>6.12</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>AVG TTR Low Frequency</td>
<td>0.5973</td>
<td>0.605</td>
<td>0.7</td>
<td>0.4889</td>
</tr>
<tr>
<td></td>
<td>372127.7565</td>
<td>340782.7435</td>
<td>2.345</td>
<td>0.028</td>
</tr>
</tbody>
</table>

A path analysis was used to analyze research question two. Standard error (S.E), critical ratio (C.R) and p-values were used to measure regression weights. Results show that C.R was significant when comparing K-Prep Math to MLU (C.R=5.624, 3.390) as well as when comparing K-Prep writing to AVG TTR (C.R=2.315). This shows that there is a significant correlation within those paths. C.R also shows a significant correlation between STAR Reading and MLU (C.R=2.334, 4.155) and STAR Math and AVG TTR (C.R=6.084). This reflects a significant correlation among these paths. P-values show a significant correlation along the following paths:

- K-Prep Math and MLU
- K-Prep Math and AVG TTR
- K-Prep Social Studies and MLU
- K-Prep Social Studies and AVG TTR
- K-Prep Writing and MLU
- K-Prep Writing and AVG TTR
- STAR Reading and MLU
- STAR Reading and AVG TTR
- STAR Math and MLU
- STAR Math and AVG TTR.
(Table 3 below).

**Figure 2. Path Analysis.**

**Table 3

Path Analysis Results**

<table>
<thead>
<tr>
<th>Traditional Measures</th>
<th>Traditional Assessments</th>
<th>Estimate</th>
<th>S.E.</th>
<th>C.R.</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLU_ST_Prompted</td>
<td>KPREP_R</td>
<td>-0.34</td>
<td>0.39</td>
<td>0.872</td>
<td>0.383</td>
</tr>
<tr>
<td>MLU_UL_Prompted</td>
<td>KPREP_R</td>
<td>-0.434</td>
<td>0.531</td>
<td>0.818</td>
<td>0.413</td>
</tr>
<tr>
<td>AVGTTR_Prompted</td>
<td>KPREP_R</td>
<td>0.007</td>
<td>0.009</td>
<td>0.756</td>
<td>0.45</td>
</tr>
<tr>
<td>MLU_ST_Prompted</td>
<td>KPREP_M</td>
<td>2.481</td>
<td>0.441</td>
<td>5.624</td>
<td>***</td>
</tr>
<tr>
<td>MLU_UL_Prompted</td>
<td>KPREP_M</td>
<td>2.038</td>
<td>0.601</td>
<td>3.39</td>
<td>***</td>
</tr>
<tr>
<td>AVGTTR_Prompted</td>
<td>KPREP_M</td>
<td>-0.044</td>
<td>0.01</td>
<td>4.207</td>
<td>***</td>
</tr>
<tr>
<td>MLU_ST_Prompted</td>
<td>KPREP_SS</td>
<td>-1.781</td>
<td>0.706</td>
<td>2.522</td>
<td>0.012</td>
</tr>
<tr>
<td>MLU_UL_Prompted</td>
<td>KPREP_SS</td>
<td>-3.396</td>
<td>0.962</td>
<td>-3.53</td>
<td>***</td>
</tr>
<tr>
<td>AVGTTR_Prompted</td>
<td>KPREP_SS</td>
<td>0.023</td>
<td>0.017</td>
<td>1.383</td>
<td>0.167</td>
</tr>
<tr>
<td>MLU_ST_Prompted</td>
<td>KPREP_W</td>
<td>-0.571</td>
<td>0.388</td>
<td>1.473</td>
<td>0.141</td>
</tr>
<tr>
<td>MLU_UL_Prompted</td>
<td>KPREP_W</td>
<td>-1.289</td>
<td>0.528</td>
<td>-2.44</td>
<td>0.015</td>
</tr>
<tr>
<td>AVGTTR_Prompted</td>
<td>KPREP_W</td>
<td>0.021</td>
<td>0.009</td>
<td>2.315</td>
<td>0.021</td>
</tr>
</tbody>
</table>
Discussion

The purpose of this study was to determine if the DM-STEAM had the potential to elicit low frequency vocabulary in comparison to the DM in spoken language samples in early adolescents with typically developing language skills. Further, the study compared standardized school assessments to the results of the DM-STEAM to determine correlation. The rationale for this research is based on the need for measurements that specifically elicit academic language for the adolescent population. Due to the lack of procedure for assessing adolescent academic language use, this population is at risk for unidentified language limitations.

In this study, researchers defined low frequency vocabulary as words used less frequently in everyday conversation and in academic contexts. Results revealed two significant findings. When comparing prompted language samples (DM-STEAM) to unprompted (DM), the DM-STEAM elicited more low frequency vocabulary. For establishing a correlation between school standardized assessments and DM-STEAM results, a path analysis showed that there was significant correlation between the two.

**Average type token ratio (AVG TTR).** This is a traditional measure of lexical sophistication which accounts for the size of the sample. Results show that the DM-
STEAM ($M = .5973$) and DM ($M = .6050$) do not show a significant difference. This means that whether the child was prompted to stay within a specific area or the child was given the freedom to use any vocabulary, their AVG TTR was similar. This means children used about the same amount of different words in both samples compared to total words.

**Mean Length of Utterance (MLU).** Comparison of the DM-STEAM to the DM revealed a significant difference ($p<.001$) regarding MLU. For this measure, the sample was analyzed for MLU. Both DM-STEAM and DM included a first impressions round, a description round, and a story telling round. For DM-STEAM, only the story portion was analyzed and the data was used to establish correlation between standardized assessments and results. The difference in MLU length with the DM-STEAM being higher may be attributed to the familiarity of the adolescent population with academic language. Further, these students were evaluated in an academic setting during the school day. Therefore, use of academic language was not a divergence from their normal setting. In contrast, some children may have needed to increase their utterances to adequately express their meaning when using academic language.

**Low Frequency.** Low frequency vocabulary is conceptualized as vocabulary used less frequently in everyday conversation. For the SUBTLEXus, a lower number reflects a sample that used more low frequency vocabulary. The DM-STEAM mean was 372127.7565. The DM had a mean of 340782.7435. For the purposes of this study, children used less low frequency vocabulary when specifically prompted compared to the DM which did not prompt for academic/subject specific vocabulary. Even so, the DM-STEAM was able to elicit vocabulary occurring at a lower frequency in the English
language with a score of 372127.7565 indicating children were still able to use more specialized vocabulary within the constraints of academic contexts.

**School Standardized Assessments**

K-PREP scores were obtained from the previous fifth grade school year. K-Prep sections for fifth grade are reading, math, social studies, and on-demand writing. There was a significant correlation between K-Prep Math and three measurements: MLU for the story portion, MLU for the whole sample, and AVG TTR. Correlations for all three measurements were significant at $p=<.001$. This correlation may be attributed to the alignment of math content represented in the math section of K-Prep and themed math knowledge. For example, themed categories such as science, engineering, and math may be closely related to mathematics curriculum knowledge.

There was also a significant correlation between K-Prep Social Studies and two measurements: MLU for the story portion ($p=.012$) and MLU for the whole sample ($p=<.001$). Likewise, this positive correlation could relate to the arts section of the STEAM categories. For K-Prep writing, there was a significant correlation between MLU for the whole sample ($p=.015$) and AVG TTR ($p=.021$) but not for the story portion. These mixed results are likely due to K-Prep writing integrating skills across the academic curriculum.

STAR Reading and Math scores were obtained from the sixth grade school year. There was a significant correlation between both STAR Reading and Math, MLU for whole sample, MLU for story portion, and AVG TTR. The results showed that high STAR reading scores were associated with lower AVG TTR and higher MLU. It also showed that high STAR Math scores were related to lower MLU and higher AVG TTR.
This may indicate that students who are good readers tend to lengthen their utterances and repeat words. Whereas, students who score higher in math tend to be more efficient with their words, resulting in more diversity in their samples.

**Limitations**

The sample size was limited due to recruitment from only one school. The sample was also limited based on the necessity of data excluded from analysis due to poor sound quality. Three student samples were also excluded based on inclusion criteria. These students did not meet the study’s standard on the TILLS SLS screener, which identified at risk students for language delays. Another limitation of this study was lack of information regarding daily academic performance (grades). If this information had been accessible, it could have been included in the overall analysis yielding a better reflection of how individual student skill correlates with the results of the DM-STEAM. Finally, the inability to give the same students the DM and the DM-STEAM is a limitation. This was not possible due to the potential influence on results because of repeated exposure to testing materials, such as Dixit cards.

**Future Research**

Additional studies with larger sample sizes are needed to confirm the preliminary results of this study. While the small sample size limited the ability to draw substantial conclusions, the results revealed a strong enough difference between the DM-STEAM and the DM to justify further research. Future research may include replicating the study with an entire sixth grade class, either in one building or across a school district. Future research may also include administering the DM-STEAM and DM to students in the same school. Future studies may also utilize this sampling method with other student
populations, such as those with specific language impairments. This could give researchers information on how to best serve adolescents with specific language impairments in terms of their academic success. Research is also on going to help establish a vocabulary list for science, technology, engineering, arts, and math to help assist educators using academic language in their lessons.

**Conclusion**

The evidence indicates that the DM-STEAM was able to elicit low frequency vocabulary compared to the DM. The DM-STEAM may be an effective language sampling procedure for eliciting lexically sophisticated language samples in the early adolescent population. Further, evidence shows that the DM-STEAM was reflective of current academic performance when compared to school standardized assessments. While future research is necessary to further evaluate and examine these claims, there is a need for a language sampling procedure that specifically elicits academic language.
References

AccessSTEM. (2007). Building capacity to include students with disabilities in science, technology, engineering, and mathematics fields. Seattle, WA: University of Washington.


Crossley, S. A., & McNamara, D. S. (2012). Predicting second language writing proficiency:


Nippold, M. A., Frantz-Kaspar, M. W., Crammond, P. M., Hayward-Mayhew, C., &


Smith, T. (2018), Thesis: The Dixit Method of Language Sampling in Early Adolescence, B.S. in Communication Disorders, Western Kentucky University, Bowling Green, KY.


Presentation, Lexington, KY.

Spencer, S., Clegg, J., & Stackhouse, J. (2010). “I don’t come out with big words like other people”: Interviewing adolescents as part of communication profiling. *Child Language Teaching & Therapy, 26*(2), 144–162.


APPENDIX A: EXAMPLES OF ILLUSTRATED DIXIT CARDS