The influence of shyness and gender on vocabulary IQ in adolescents

Traynor, Julie

Universiteit Leiden

Psychologie
Faculteit der Sociale Wetenschappen

The influence of shyness and gender on vocabulary IQ in adolescents

Julie Traynor

Master Thesis Child & Adolescent Psychology

Date: 08/12/2025

Supervisor: Dr Ivan Simpson-Kent

Second reader: Dr Kelly Ziemer

Abstract

Shyness is an unease towards social interaction and evaluation and is present in up to 60% of adolescents, with females indicating greater shyness. Shy children were scored as having worse language abilities by teachers compared to their non-shy peers. This study aimed to examine the relationship between shyness and vocabulary IQ amongst a general population of 4160 adolescents aged 13-24. The open-source data from The Vocabulary IQ test and questions S2 and S4 from the supplementary questionnaire were assessed in order to assess the research question. The research hypothesised that higher shyness composite scores would be associated with lower scores on the Vocabulary IQ test, that gender moderates this association, and that the relationship is stronger amongst female participants. This study's sample comprised of 42.67% males and 52.14% females. Although the multiple linear regression model was significant, results from analysis showed no significant relationship between shyness and vocabulary when accounting for gender, age, and English as a native language. Additionally, gender did not significantly moderate any association between shyness and vocabulary. However, age and English as a native language were significantly correlated with vocabulary IQ scores, with older respondents and those with native English scoring higher. Results of the current study also suggested a non-significant positive relationship between shyness and vocabulary IQ. The absence of an assessment involving a social aspect of vocabulary evaluation in this study may have influenced shy adolescents' performance. These findings have implications for teacher perceptions of shy children's language abilities and add to the research involving measures without social interaction to assess shy adolescents' vocabularies.

Laymann's Abstract

Past research on shy young peoples' language skills has used test measures that favour more social personalities. Results from these studies have suggested that shy young people have lower vocabulary abilities. There is less research on vocabulary knowledge in shy adolescents, so this study focused on this group. The current study used a method of studying vocabulary that did not contain any social interaction: The Vocabulary IQ test. This study focused on the 4160 adolescents aged 13-24 years old. Answers from two additional questions were combined to determine the adolescent's level of shyness: "I avoid contact with others" and "I'd rather not people pay attention to me". Statistical analyses indicated that levels of shyness were not related to vocabulary knowledge and that there were no differences between male and female adolescents. This contrasts with previous studies that found that higher shyness or introversion was associated with lower language skills. This study found that age and English as a native language were related to vocabulary ability, with older adolescents and those with native English scoring better on the vocabulary test. It is recommended that future studies examining shy adolescents use an assessment which does not involve conversation or speaking. Additionally, future research would benefit from including socio-economic status information when assessing vocabulary knowledge, something this current study did not have access to. Overall, this

research adds to the literature by showing no relationship between shy adolescents and their vocabulary knowledge abilities.

Introduction

Shyness describes an individual's unease in social interactions and in perceived social evaluation (Rubin, Coplan, and Bowker, 2009). Shyness involves an approach-avoidance conflict, whereby people crave social interaction, but their fear and anxiety trigger social avoidance (Coplan et al., 2004). It is suggested that up to 60% of secondary school students report themselves as shy (Onukwufor & Iruloh, 2017). Shyness and social avoidance have consistently been associated with internalising problems during childhood and adolescence (Muris et al., 2003; Sette et al., 2023). Internalising problems are inwardly focused emotional issues, including anxiety, depression, loneliness, and trauma-related issues (Achenbach & Edelbrock, 1978). In children ages 7-11 years old, social avoidance has been linked to loneliness and depression (Sette et al., 2023). Similarly, self-reported shyness in 9-12 year old children was related to higher social anxiety and depression symptoms (Coplan et al., 2013). These internalising problems have been suggested to be more prevalent in withdrawn males compared to withdrawn females (Coplan & Weeks, 2009), indicating differences based on gender in shy adolescents.

Gender differences in shyness have been reported in various studies. Rather than remaining stable over time, shyness is thought to increase into adolescence and has been noted to vary according to gender (Kaveworld et al., 2012). Doey et al., (2014) suggest in their study on shyness and gender differences that shyness has a greater developmental risk for boys compared to girls. Despite this assertion, female adolescents are noted to have greater shyness than their male counterparts, with the level of shyness in girls significantly higher than that of boys aged between 10 and 23 years old (Wang et al., 2020). Additionally, 76% of female students aged 13-14 years old report being shy compared to 44% of male students (Onukwufor & Iruloh, 2017). Therefore, differences in shyness during adolescence may be influenced by gender, with female adolescents reporting higher shyness.

Vocabulary is defined by Random House Webster's Unabridged Dictionary (Flexner, 1997) as the "stock of words used by or known to a particular people or group of persons". Vocabulary is a core factor of language and a significant indicator of overall verbal intelligence in children aged 6-14 and young adults aged 18-25 (Segbers & Schroeder, 2017). Vocabulary knowledge has shown to be a more powerful predictor of language compared to verbalisation (Vallotton & Ayoub, 2011). Ahmed et al., (2016) also suggest that vocabulary knowledge has one of the strongest links to reading performance in 7th-12th graders (ages 12-18 year old). Vocabulary has been shown to improve with age. Age was noted by Keuleers et al., (2015) as the most important predictor of vocabulary abilities across the lifespan. Additionally, native speakers have been shown to have greater vocabulary than non-native or bilingual language speakers in a given language (Bialystok & Luk, 2012).

Shyness in children and adolescents has been connected to language capabilities. Spere et al. (2004) found that shy children as young as preschool age performed worse in receptive and expressive

measures of language, including vocabulary knowledge, compared to their non-shy peers. Coplan and Weeks (2009) found that children aged 6-7 years old with higher levels of shyness had lower pragmatic language skills. This study design involved children reading aloud and completing stories using pragmatic language abilities (using context clues to understand meaning and communicate with others). A small gender effect which did not reach significance was noted in this research. Female participants who scored lower in pragmatic language skills frequently also had greater levels of social withdrawal. Additionally, girls scored significantly higher than boys in the social anxiety subscale. An association between shyness and language may therefore be influenced by gender.

The connection between higher shyness and language problems can further develop into late adolescence and early adulthood (Durkin et al., 2017). Despite this, minimal research has been conducted on this relationship in adolescents. Samadi et al., (2013) found that shy adolescent students aged 14-18 years old scored differently on reading comprehension, with a tendency for higher reading scores. However, this study did not clearly establish a positive or negative association between shyness and reading comprehension. To further understand this association, it is necessary to look to research on adolescents with a specific language impairment (SLI). An association between those with an SLI and increased shyness has been noted, with a group of 16–17 year-olds with SLI demonstrating significantly higher shyness compared to a peer group with typical language abilities (Wadman et al., 2008). Similarly, young adults between ages 17-24 years old with a language impairment scored higher in shyness compared to their age matched peers with typical language abilities (Durkin et al., 2017). Young people with an SLI have notably poorer receptive vocabulary (Hafeez et al., 2023) and express an increased deficiency in vocabulary ability as they age, up to 21 years old (Rice & Hoffman, 2015). Accordingly, a link has been noted between language ability (including vocabulary) and shyness in adolescents with a specific language impairment, with those who show lower vocabulary abilities having higher levels of shyness.

Despite findings in adolescents with SLIs, there are mixed explanations for the association between shyness and language abilities in typically developing children. Coplan et al., (2013) suggest that shy children speak less and therefore have less opportunity to develop and practise language skills, leading to comparatively lower linguistic abilities. Contrastingly, Zhu et al. (2019) propose that this association may instead reflect difficulties with performance due to the social nature of tests, rather than a true limitation in language competence. For example, a commonly used test of vocabulary and language abilities in children is The Peabody Picture Vocabulary Test (PPVT-4; Maddux, 1999), which was utilised by the aforementioned studies by Spere et al. (2004) and Zhu et al. (2019). The PPVT involves a level of social interaction, as the interviewer reads a word aloud, followed by the participant choosing an illustration to best fit the word (Rowe et al., 2012). Therefore, the use of social interaction in language tests may better explain a negative association between shyness and language abilities in children. Findings may be related to a resistance by shy individuals

to respond rather than a deficit in language (Smith Watts et al., 2014). Anugrah and Idayani's (2024) study utilised Google Forms, a method of data collection which involved no social interaction. Their results showed no significant correlation between introversion and reading comprehension. Additionally, Hughes and Coplan (2010) found that teachers rated shy children's reading skills as poorer compared to their peers, but no association between shyness and reading comprehension was suggested in standardised tests. They therefore suggested that shyness inhibits the display of certain academic abilities in specific contexts, but it is not directly related to academic knowledge or skills (Hughes & Coplan, 2010). The association between shyness and reading abilities (such as vocabulary) may therefore vary depending on the measure used to assess reading abilities.

The present study

Research has noted a link between vocabulary knowledge and shyness in young children and adolescents with SLI (Spere et al., 2004; Durkin et al., 2017). Samadi et al., 2013 showed a correlation between shyness and reading comprehension in adolescents, but did not clearly establish a negative or positive association. The current study therefore aims to investigate the association between shyness and vocabulary utilising a regression analysis. Previous research on vocabulary in shy young people has utilised measures which involve a social-performance aspect of assessment, such as answering a question out loud (Pickering et al., 2022). It has been suggested that this may be influencing results which show that shy children perform worse on tests of vocabulary (Zhu et al., 2019). Therefore, this study analysed adolescents' vocabulary IQ without this social-performance aspect. A measure of vocabulary and shyness which does not involve social interaction–the Vocabulary IQ test– was implemented for this study. The present study aimed examine the relationship between shyness and vocabulary knowledge in adolescence. This study also examined the moderating effect of gender, as previous research has found that adolescent females often show higher shyness compared to their male peers (Wang et al., 2020).

The current study is a relevant addition to the literature on shyness and language abilities in adolescents. Findings suggested that shy adolescents may have a resistance to respond in a testing environment rather than a true deficit in language (Smith Watts et al., 2014). This also impacted educational settings, as teachers have been found to underestimate shy children's reading ability (Hughes and Coplan, 2010). Perceived negative evaluation from teachers has been associated with higher internalising problems in children (Jellesma et al., 2015). This may be particularly impactful, as shy adolescents are reported to generally have higher internalising problems such as loneliness, anxiety, and depression (Coplan et al., 2013). As vocabulary is a core factor of language in adolescence and young adults (Segbers & Schroeder, 2017), assessing vocabulary abilities in shy adolescents may aid in understanding their general language abilities. Undergoing this analysis

utilising a test which does not involve any social interaction aspects may allow for more accurate vocabulary scores from shy adolescents.

Hypotheses

The first aim of this study was to examine if a significant negative relationship between shyness and vocabulary scores in adolescents (ages 13 to 24 years old) can be seen using the Vocabulary IQ test. More specifically, this study predicted (H1) that higher levels of shyness are associated with lower vocabulary IQ scores. This hypothesis was grounded in previous findings whereby lower vocabulary knowledge was associated with difficulties such as increased shyness (Spere et al., 2004; Thornton, 2022). This hypothesis utilised a measure of both vocabulary and shyness which does not include a social aspect: The Vocabulary IQ test. The use of this test reduces the possibility that underperformance by shy adolescents is due to social test demands (Zhu et al., 2019). The second goal of this study was to examine the role of gender in the interaction between vocabulary and shyness. It was hypothesised (H2) that gender moderates the relationship between shyness and vocabulary, with the strength of this relationship differing between male and female adolescents. This hypothesis was based on previous research suggesting a small, gendered effect between language ability and shyness in younger children, and notable differences in shyness between male and female adolescents (Coplan & Weeks, 2009; Wang et al., 2020). Following the evaluation of the role of gender, it was predicted that (H3) female respondents experience a stronger association between shyness and vocabulary than male respondents. This hypothesis draws upon research by Doey et al., (2014) and Wang et al., (2020), who both concluded that female adolescents experienced higher levels of shyness compared to their male peers.

Methods

Design

This study involved a between-subjects cross-sectional design. Data were collected from July 2017 to March 2018, and were available from an open-source database, the Open-Source Psychometrics Project (www.openpsychometrics.org). This study was a correlational design, evaluating the association of the independent variable of shyness on the dependent variable of Vocabulary IQ scores. The between-subjects factor of gender was used as a moderator to examine the relationship between shyness and vocabulary. Vocabulary knowledge was measured using the Vocabulary IQ test, while shyness was measured using questions from the Vocabulary IQ supplementary survey. Age and English as a native language were included in this design as control variables.

Respondents

Data were obtained from an open-source database, the Open-Source Psychometrics Project (www.openpsychometrics.org). Limited information on recruitment was given regarding this sample. The total sample size of this dataset was 12173. However, for the purpose of this study, data were filtered to only include adolescent respondents. The age range in this study was between 10-24 years, which aligns with modern growth expectations and social theories regarding adolescence, as outlined by Sawyer et al. (2018). As the youngest available age in this dataset was 13 years old, the finalised age range for the current study's respondents was from 13-24 years old (M =19.26 years old, SD = 2.97 years). Following the exclusion of respondents over 24, the sample size for this study was 4396. Regarding gender, 42.67% of respondents identified as male (n = 1875), 52.14% identified as female (n = 2291), 3.78% identified as "other" (n = 166), and 1.46% did not answer (n = 64). To focus on male and female differences, "other" and those who did not answer were removed from the dataset. This aligns with previous research, which only focused on female-to-male differences rather than all gendered group differences (i.e Wadmann et al., 2008). After these gender groups were excluded, participants over 24 were removed, and those who did not complete the test or the supplementary questionnaire were removed; the final total sample was n = 4160. Of this sample, 69% of respondents had English as their native language (n = 2874).

Measures

The Vocabulary IQ test

The dependent variable of vocabulary knowledge was measured using the Vocabulary IQ test. The Vocabulary IQ test (https://openpsychometrics.org/tests/VIQT/) consisted of 45 questions, each containing five words for respondents to choose from. Respondents identified two words from the list of five which had the same meaning. A -0.35 point penalty was given for incorrect answers. Full

vocabulary scores were calculated by (number of correct answers) - 0.35(number of incorrect answers). Lower scores on this test indicated worse vocabulary ability. The independent variable of shyness was measured using two questions from the 35-question supplemental personality items of the Vocabulary IQ test: "S2: I avoid contact with others", and "S4: I'd rather not people pay attention to me". Although the initial proposal included another question in this measure, "S24: I emerge as a natural leader", it was removed. Cronbach's alpha when the question S24 was included with S2 and S4 was $\alpha = 0.13$, indicating poor internal consistency. The remaining two questions were scored on a five-point Likert scale (1 = Disagree, 5 = Agree). The resulting scores from each question were summed into a single combined composite score measure of shyness. Larger composite scores (maximum of 10) indicated greater levels of shyness. Age and gender were measured using a supplementary questionnaire before the test began by asking respondents to fill out their age and to choose their gender from the options 'Male', 'Female' and 'Other'.

Procedure

Ethical approval for this study was obtained from the Leiden University Psychology Research Ethics Committee (CEP number and approval date: SimpsonKent-V1-5058, 2023-10-29-I.L.). The data were collected online between July 2017 and March 2018. Data were accessed through an open-source repository from the Open-Source Psychometrics Project data page (https://openpsychometrics.org/_rawdata/). Respondents were informed at the start of the test that their responses would be used for research and were asked to confirm that their answers were accurate. No consent form was signed by respondents. However, following the vocabulary questions, respondents were asked if they made their best attempt and if their data can be used for research and analysis. Respondents were informed first on an opening screen of what the test entailed before beginning. On this opening page, respondents were informed of the task and instructed not to guess due to a point penalty for incorrect answers, and instead to hit "I don't know". Respondents then completed the 45 vocabulary questions. They were then asked to complete a 3–5-minute supplemental survey with 35 personality and demographic questions. At the end of the test, respondents were shown their Vocabulary IQ score and informed of the normative scores of the population. All data were anonymous.

Statistical Analyses

Statistical analyses were undertaken using the R programming language with R-Studio version 2023.12.1+402 (RStudio Team, 2023). The internal consistency of the question items was assessed using Cronbach's alpha, with items being considered suitable if the internal consistency was $a>0.6$. The relationship between each of the items was determined using a Pearson's Correlation, and items were deemed acceptably correlated when $r>0.4$.

The association between shyness and vocabulary knowledge was assessed using a linear regression analysis, with a p-value of <.05 considered statistically significant. The assumptions of regression were checked prior to completion of the linear regression. Linearity was checked by plotting the residuals versus fitted errors. Independence of residual errors was checked using the Durbin-Watson test, with independence being accepted if results are between 1.5 and 2.5 (Turner, 2020). To test for normality of residuals, a Q-Q plot was analysed, with normality accepted if residuals approximately follow a straight line. The Variance Inflation Factor (VIF) was used to test for multicollinearity, with results of 1 indicating no multicollinearity and <5 indicating only moderate, acceptable multicollinearity (Singh, 2024). The moderating effect of gender on this association was determined with a moderation regression analysis. Separate regression analyses were initially planned to test the hypothesis that female adolescents experience a stronger correlation between shyness and vocabulary compared to male adolescents. However, these analyses were not conducted following the results of the initial moderation analyses, as noted in the 'Regression Analyses' results section. Finally, The Benjamini-Hochberg (1995) procedure was used to control the False Discovery rate, which aided in reducing the proportion of false positives while maintaining high statistical power (Guo and Rao, 2008).

Results

A multiple linear regression was conducted to examine the impact of shyness as a predictor of vocabulary knowledge and whether this relationship was moderated by gender. Both age of respondents and English as a native language were included in the analyses as control variables. The moderating effect of gender was assessed by the interaction between shyness and gender on vocabulary in the regression model. Age was used as a control variable in this study to account for potential developmental differences in vocabulary knowledge and shyness (Durkin et al., 2017). English as a native language was an additional control variable, as non-native English speakers may have less proficiency in English vocabulary (Lessard-Clouston, 2006). The median shyness score of all respondents was 6. The lowest Vocabulary IQ score was -15.75, with the highest being 45.00. The mean Vocabulary IQ score was 25.65. (*SD* = 9.739).

Assumptions of regression

Assumption tests of a linear regression were conducted. A scatterplot of residuals versus fitted values assessed linearity. Visual results showed a linear relationship between shyness and vocabulary, and that residual errors had a mean value of zero (Appendix A). To test for independence of residual errors, the Durbin-Watson Test was performed. The test presented a result of 1.829 which falls in an acceptable range (1.5 < 2.5), indicating that the residuals are independent. A Q-Q plot was utilised to test for normal distribution of residuals. Most residuals aligned with the expected diagonal line, with some deviation at the tails (Appendix B). These deviations suggest a small variation from normality; however, the assumption of normality was accepted due to the large sample size. Finally, no multicollinearity was indicated, as the VIF for all factors was 1.00. As the assumptions of regression were met, a regression analysis was conducted.

Regression analysis

A linear regression was conducted to examine Hypothesis 1, that shyness significantly negatively predicts vocabulary knowledge in adolescents. The overall regression model was statistically significant $R^2 = .21$, $F(5, 4160) = 221.20$, $p < .001$. Shyness was a non-significant positive predictor of Vocabulary IQ score $B = .16$, $t(4160) = 1.81$, $SE = .09$, $p = .071$ 95% CI [-.01, .032] when age and English as a native language were included as control variables. Only 21% of variance in vocabulary was explained by shyness and gender. There was a statistically significant relationship between vocabulary and age $B = .87$, $SE = .29$, $t(4160) = 19.04$, $p = .003$ 95% CI[0.78, 0.96]. This indicated that vocabulary and age were positively associated with each other, with older adolescents scoring higher in vocabulary IQ scores. Native English was also a strong positive predictor of vocabulary IQ score, with native speakers scoring higher than non-native speakers $B = -7.44$, $SE = 0.29$, $t(4160) = -25.84$, $p < .001$. However, in contrast to Hypothesis 1, there was a non-significant

positive rather than negative relationship between shyness and vocabulary, with higher shyness scores related to higher vocabulary scores. Therefore, Hypothesis 1 is not supported.

Another linear regression model was conducted to examine if gender moderated the relationship between vocabulary and shyness. Overall, the model was statistically significant $R^2 = .21$, $F(5, 4160) = 221.2$, $p < .001$. However, the interaction between vocabulary and gender with shyness was not statistically significant $B = .16$, $t(4160) = 1.23$, $SE = .12$, $p = .071$, 95% CI [-0.09, .038]. This indicated that gender did not moderate the relationship between vocabulary IQ and shyness, and therefore that this relationship does not differ significantly between female and male adolescents. Only a small amount of the variance in shyness (21%) was explained by the interaction between vocabulary and gender. Hypothesis 2, that gender significantly moderates the relationship between shyness and vocabulary was therefore not supported.

Additionally, as gender did not appear to moderate the relationship between vocabulary and shyness, no additional analysis was undertaken for the third hypothesis which stated that female adolescents had a stronger association between vocabulary and shyness. The previous moderation analysis revealed that the strength of the relationship between shyness and vocabulary did not differ based on the gender of the respondents.

Discussion

This study investigated the relationship between shyness and vocabulary knowledge in adolescents aged 13-24 years old. Previous research on personality traits' influence on reading comprehension, including vocabulary, is limited (Krach et al., 2015). This study aimed to fill this gap in literature by examining the association between shyness and vocabulary with gender as a moderator. Including gender as a moderator was based on findings from Doey et al., (2014) and Wang et al., (2020), which suggested that female adolescents experienced higher levels of shyness compared to their male peers. Results of the current study showed a non-significant positive relationship between shyness and vocabulary IQ, such that higher scores in shyness related to higher scores on the vocabulary IQ test. Moreover, no significant interaction between vocabulary IQ, shyness, and gender was found. This study therefore indicated that there is no difference in the aforementioned association between shyness and vocabulary between male and female adolescents.

Previous studies found that shy preschool children performed worse in vocabulary tasks compared to their non-shy peers (Spere et al., 2004). Similarly, Durkin et al.'s (2017) study of adolescents with language impairments showed that they scored higher in shyness compared to their age-matched peers without any language difficulties. In the current study, the overall regression model was statistically significant. However, only a small amount of variance in vocabulary (21%) was explained by shyness and gender. Therefore, the statistical significance of the model may have been a result of the large sample size ($N = 4160$). Prior research suggested that shy children and adolescents perform worse in tests of language and vocabulary, and that young people with language difficulties exhibit higher levels of shyness (Zhu et al., 2019; Wadman et al., 2008). These earlier studies indicate a negative association between vocabulary and shyness. However, the current study identified a non-significant positive correlation, whereby shy adolescents displayed higher vocabulary IQ scores.

This disparity in findings may reflect the method of vocabulary knowledge assessment used. Previous studies, such as that by Spere et al., (2004), assessed vocabulary and language abilities using the Peabody Picture Vocabulary Test (PPVT-4; Maddux, 1999). The PPVT involves a verbal social interaction between the interviewer and interviewee (Rowe et al., 2012). Shy children may have therefore performed worse on these tests due to a resistance to respond to the social situation rather than a deficit in vocabulary knowledge (Smith Watts et al., 2014). This phenomenon is also supported by findings by Hugh and Coplan (2010), who suggested that shyness inhibits the display of academic abilities in certain contexts, but it is not related to academic skills in practice. A strength of the present research is therefore the use of a vocabulary measure that does not involve a social interaction component: The Vocabulary IQ test. This methodological difference in vocabulary assessment may explain why the current study did not find any significant relationship between shyness and vocabulary IQ. This study also supports the findings suggesting that shy adolescents do not have

worse language abilities, but rather that social test demands impeded their results. This study alternatively suggested that shy adolescents may have better vocabulary abilities, although these results did not reach significance. Similarly, Anugrah and Idayani (2024) found no significant correlation between introversion and reading comprehension. Their study also administered a language test involving no social interaction. It is therefore recommended that future research continue to use tests involving less social interaction when investigating shyness and language abilities such as vocabulary.

Previous studies have suggested that shyness functionally impacts males and females differently. Doey et al., (2014) suggest that shyness poses as a greater developmental risk for boys compared to girls. Alternatively, girls between the ages of 10 and 23 were notably shyer than their male counterparts (Wang et al., 2020; Onukwufor & Iruloh, 2017). These distinctions in shyness between males and females indicated that there could be a variation in vocabulary results based on gendered shyness. This was also suggested by Coplan and Weeks (2009) in their study on 6–7-year-olds. Despite previous studies' findings, no significant difference between male and female respondents in the relationship strength between shyness and vocabulary was found. Therefore, further investigation into gender, shyness, and vocabulary knowledge is needed to establish if gender does moderate an association between vocabulary and shyness. Alternatively, future studies may benefit from assessing this interaction in younger children, age matched to Coplan and Week's (2009) study.

The present analysis demonstrated a strong positive relationship between age and vocabulary knowledge. This aligns with previous research on vocabulary, finding that vocabulary breadth increases with age. One study by Keuleers et al., (2015) determined that 17% of its population's variance in vocabulary was explained by older age. Additionally, the current findings showed that native English status was positively associated with vocabulary IQ score, with non-native speakers scoring lower in vocabulary. Previous findings also suggest that native speakers have a larger vocabulary than non-native or bilingual speakers, particularly in adolescent and adult speakers (Bialystok & Luk, 2012). A strength of the current study, therefore, was accounting for native English status and age as control variables in the analysis. Future studies may benefit from accounting for or removing non-native speakers from their analyses to reduce confounding, as was done by Keuleers et al. (2015), in addition to including age as a control variable.

Current results indicate that only a small amount of variance in vocabulary IQ is accounted for by shyness and covariates. This suggests that an unassessed factor or factors may explain additional variance in vocabulary IQ scores. Socioeconomic status (SES) has been shown to predict differences in children's language development. According to Schwab and Lew-Williams (2016), those at the lower end of the SES spectrum received less high-quality language experience, which

impacted vocabulary development. Parental SES and parent education were large predictors of vocabulary knowledge (Thornton, 2022). Norbom et al., (2022) found that youth between 3-21 years old from lower SES families had poorer language related abilities including vocabulary knowledge compared to their age matched peers. A potential limitation of the present study, therefore, is the absence of SES data. SES may explain part of the variance associated with vocabulary, potentially altering the size or direction of the relationship. Future studies assessing vocabulary IQ should benefit from including supplementary questions on parental education and SES status.

An additional limitation in the methodology of this study may also explain its divergent findings. The variable of 'shyness' was comprised of two questions from the Vocabulary IQ supplementary questionnaire: "S2: I avoid contact with others", and "S4: I'd rather not people pay attention to me". Although these questions target the social-avoidance paradigm associated with shyness (Coplan et al., 2004), they do not fully address the in-the-moment unease in social interactions (Rubin, Coplan, and Bowker, 2009). The full aspects of shyness may not have been accurately assessed using this current measure. Future studies may benefit from using a broader measure of shyness or including further supplementary questions based on a more robust measure of shyness. One such measure is the Revised Cheek and Buss Shyness Scale (RCBS; Cheel, 1983), a 13-item scale which assesses discomfort and inhibitions in the presence of other people.

This studies' findings have practical implications for adolescents, particularly those in educational settings. Notably, Hughe and Coplan's (2010) study on teachers' perceptions of academic performance revealed that they viewed shy young people as less intelligent than their non-shy peers. Shy youth were rated worse in reading skills by teachers, but no such relationship appeared in standardised tests. In contrast, the current study reveals no significant relationship between shyness and vocabulary knowledge. These findings support the claim that shy adolescents are no less intelligent than their non-shy peers. Although this study focused on vocabulary knowledge rather than overall language ability, vocabulary was previously noted as a significant indicator of overall verbal intelligence in children aged 6-14 and young adults aged 18-25 (Segbers & Schroeder, 2017). Vocabulary knowledge has also been linked to reading comprehension in 7^{th}-12^{th} graders aged 12 to 18 years old (Ahmed et al., 2016). Although relying solely on supplementary Vocabulary IQ questions to measure shyness was a limitation, the overall use of the Vocabulary IQ test was a strength of this study as vocabulary is a strong indicator of overall language abilities. Future studies may benefit from the analysis of vocabulary IQ to understand adolescents' language abilities. These findings may also aid in the changing of perceptions around shy children's abilities, particularly in language comprehension and intelligence.

In conclusion, the present research demonstrated a non-significant positive relationship between shyness and vocabulary knowledge, with only a small amount of variance in vocabulary explained by shyness. These findings contrast suggestions from previous studies. Additionally, this non-significant relationship did not differ based on gender. Both age and native-English ability were significantly associated with vocabulary IQ score. The small amount of variance accounted for by shyness suggests that an unstudied factor or factors may account for more variance in vocabulary IQ. This study did not ask respondents for socio-economic status (SES). SES has been shown to impact vocabulary knowledge, with those from lower SES having less advanced vocabulary (Norbom et al., 2022). This was a notable limitation of the current methodology, and future studies may benefit from including SES as a control variable when examining vocabulary knowledge. An additional possible limitation of this study was the measure of shyness used. Future studies may benefit from using a broader measure of shyness. Despite these potential limitations, a strength of this study is the implementation of a vocabulary IQ assessment that does not involve a social component. Findings suggest that the method of assessment should be carefully considered as assessment involving social interaction may influence results in shy populations. These findings have implications for the perception of shy adolescents in education. This study also supports the use of tools without social aspects in future studies examining shy adolescents.

References

Achenbach, T. M., & Edelbrock, C. S. (1978). The classification of child psychopathology: a review and analysis of empirical efforts. Psychological bulletin, 85(6), 1275.

Ahmed, Y., Francis, D. J., York, M., Fletcher, J. M., Barnes, M., & Kulesz, P. (2016). Validation of the direct and inferential mediation (DIME) model of reading comprehension in grades 7 through 12. *Contemporary Educational Psychology*, *44*, 68-82.

Anugrah, B., & Idayani, A. (2024). The Correlation between Introvert Personality and Reading Comprehension at the Sixth Semester Students of English Language Education-FKIP UIR. *EJI (English Journal of Indragiri): Studies in Education, Literature, and Linguistics*, *9*(1), 204-218

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal statistical society: series B (Methodological), 57(1), 289-300.

Bialystok, E., & Luk, G. (2012). Receptive vocabulary differences in monolingual and bilingual adults. *Bilingualism: Language and Cognition*, *15*(2), 397–401. doi:10.1017/S136672891100040X

Cheek, J. M. (1983). Revised cheek and buss shyness scale. *Journal of Personality Assessment*.

Coplan, R. J., & Weeks, M. (2009). Shy and soft-spoken: shyness, pragmatic language, and socio-emotional adjustment in early childhood. Infant and Child Development: An International Journal of Research and Practice, 18(3), 238-254.

Coplan, R. J., Prakash, K., O'Neil, K., & Armer, M. (2004). Do you" want" to play? Distinguishing between conflicted shyness and social disinterest in early childhood. Developmental psychology, 40(2), 244.

Coplan, R. J., Rose-Krasnor, L., Weeks, M., Kingsbury, A., Kingsbury, M., & Bullock, A. (2013). Alone is a crowd: social motivations, social withdrawal, and socioemotional functioning in later childhood. Developmental psychology, 49(5), 861.

Doey, L., Coplan, R. J., & Kingsbury, M. (2014). Bashful boys and coy girls: A review of gender differences in childhood shyness. *Sex roles*, *70*, 255-266.

Durkin, K., Toseeb, U., Botting, N., Pickles, A., & Conti-Ramsden, G. (2017). Social confidence in early adulthood among young people with and without a history of language impairment. *Journal of Speech, Language, and Hearing Research*, *60*(6), 1635-1647.

Flexner, S. B. (1997). Random house Webster's unabridged electronic dictionary.

Guo, W., & Rao, M. B. (2008). On optimality of the Benjamini–Hochberg procedure for the false discovery rate. Statistics & probability letters, 78(14), 2024- 2030.

Hafeez, H., Yasmin, T., Raza, M. H., Mubarak, L., Ashraf, K., Samra, M. M., & Basra, M. A. R. (2023). Receptive vocabulary, memory span, and speech articulation in Pakistani children with developmental language disorders. *Child Neuropsychology*, *29*(3), 391-412.

Hughes, K., & Coplan, R. J. (2010). Exploring processes linking shyness and academic achievement in childhood. *School Psychology Quarterly*, *25*(4), 213.

Jellesma, F. C., Zee, M., & Koomen, H. M. (2015). Children's perceptions of the relationship with the teacher: Associations with appraisals and internalizing problems in middle childhood. *Journal of Applied Developmental Psychology*, *36*, 30-38.

Karevold, E., Ystrom, E., Coplan, R. J., Sanson, A. V., & Mathiesen, K. S. (2012). A prospective longitudinal study of shyness from infancy to adolescence: Stability, age-related changes, and prediction of socio-emotional functioning. Journal of abnormal child psychology, 40, 1167-1177.

Keuleers, E., Stevens, M., Mandera, P., & Brysbaert, M. (2015). Word knowledge in the crowd: Measuring vocabulary size and word prevalence in a massive online experiment. *The Quarterly Journal of Experimental Psychology*, *68*(8), 1665–1692. https://doi-org.ucd.idm.oclc.org/10.1080/17470218.2015.1022560

Krach, S. K., McCreery, M. P., Jones, W. P., & Loe, S. A. (2015). Do Dispositional Characteristics Influence Reading?: Examining Personality and Reading Fluency. *Reading Psychology*, 470-486.

Maddux, C. D. (1999). Peabody Picture Vocabulary Test (PPVT-III).*Diagnostique*, *24*(1-4), 221–228. https://doi.org/10.1177/153450849902401-419

Muris, P., Meesters, C., & Spinder, M. (2003). Relationships between child-and parent-reported behavioural inhibition and symptoms of anxiety and depression in normal adolescents. Personality and Individual Differences, 34(5), 759-771.

Onukwufor, J. N., & Iruloh, B. R. N. (2017). Prevalence, Gender and Level of Schooling Differences in Secondary School Students Level of Shyness. *Journal of Education and Practice*, *8*(2), 93-100.

Pickering, H. E., Parsons, C., & Crewther, S. G. (2022). The effect of anxiety on working memory and language abilities in elementary schoolchildren with and without Additional Health and Developmental Needs. Frontiers in Psychology, 13, 1061212.

Rice, M. L., & Hoffman, L. (2015). Predicting vocabulary growth in children with and without specific language impairment: A longitudinal study from 2; 6 to 21 years of age. *Journal of Speech, Language, and Hearing Research*, *58*(2), 345-359.

Rowe, M. L., Raudenbush, S. W., & Goldin-Meadow, S. (2012). The pace of vocabulary growth helps predict later vocabulary skill. *Child development*, *83*(2), 508-525.

RStudio Team. (2023). RStudio: Integrated Development for R (Version 2023.06.0-421). Boston, MA: RStudio, PBC. Retrieved from https://posit.co/products/open-source/rstudio/.

Rubin, K. H., Coplan, R. J., & Bowker, J. C. (2009). Social withdrawal in childhood. Annual review of psychology, 60(1), 141-171.

Samadi, F., Maghsoudi, M., & Azizmohammadi, F. (2013). The Relationship between Shyness and Local/Global Reading Comprehension. *Language In India*, *13*(4).

Sawyer, S. M., Azzopardi, P. S., Wickremarathne, D., & Patton, G. C. (2018). The age of adolescence. The lancet child & adolescent health, 2(3), 223-228.

Schwab, J. F., & Lew-Williams, C. (2016). Language learning, socioeconomic status, and child-directed speech. *Wiley interdisciplinary reviews. Cognitive science*, *7*(4), 264–275. https://doi.org/10.1002/wcs.1393

Segbers, J., & Schroeder, S. (2017). How many words do children know? A corpus-based estimation of children's total vocabulary size. *Language Testing*, *34*(3), 297-320.

Sette, S., Zava, F., Baumgartner, E., Laghi, F., & Coplan, R. J. (2023). Examining links between social withdrawal subtypes and internalizing problems among Italian primary school children. *European Journal of Developmental Psychology*, *20*(2), 268-286.

Singh, V. (2024, November 18). *Variance Inflation Factor (VIF): Addressing Multicollinearity in Regression Analysis*. Datacamp.com; DataCamp. https://www.datacamp.com/tutorial/variance-inflation-factor

Smith Watts, A. K., Patel, D., Corley, R. P., Friedman, N. P., Hewitt, J. K., Robinson, J. L., & Rhee, S. H. (2014). Testing alternative hypotheses regarding the association between behavioral inhibition and language development in toddlerhood. *Child development*, *85*(4), 1569-1585.

Spere, K. A., Evans, M. A., Hendry, C. A., & Mansell, J. (2009). Language skills in shy and non-shy preschoolers and the effects of assessment context. *Journal of Child Language*, *36*(1), 53-71.

Thornton, E. R. (2022). Socioeconomic inequalities in vocabulary and their implications for educational attainment and mental health in adolescence. The University of Liverpool (United Kingdom).

Turner, P. (2020). Critical values for the Durbin-Watson test in large samples. *Applied Economics Letters*, *27*(18), 1495–1499. https://doi-org.ucd.idm.oclc.org/10.1080/13504851.2019.1691711

Wadman, R., Durkin, K., & Conti-Ramsden, G. (2008). Self-esteem, shyness, and sociability in adolescents with specific language impairment (SLI). Journal of speech, language, and hearing research, 51(4), 938-952.

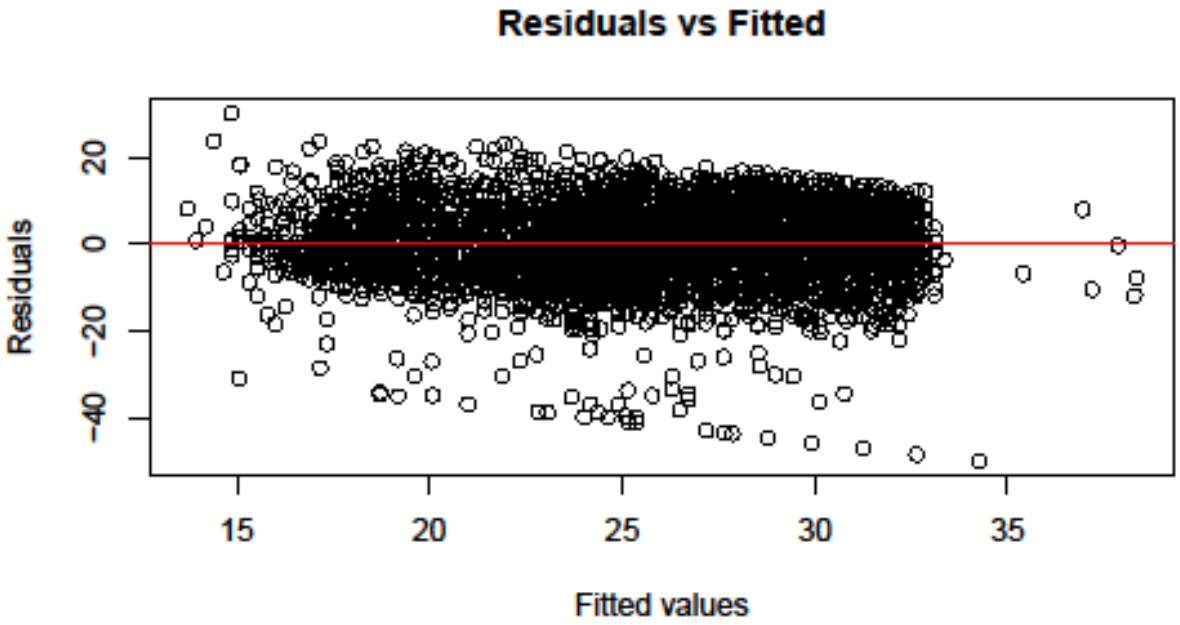
Wang, P., Yan, Y., Gao, F., Zhang, R., Wang, J., Zhan, X., & Tian, Y. (2020). The effect of shyness on adolescent network problem behavior: the role of gender and loneliness. *Frontiers in Psychology*, *11*, 803.

Xie, F., Xin, Z., Chen, X., & Zhang, L. (2019). Gender difference of Chinese high school students' math anxiety: The effects of self-esteem, test anxiety and general anxiety. *Sex roles*, *81*, 235-244.

Zhu, J., Li, Y., Wood, K. R., Coplan, R. J., & Chen, X. (2019). Shyness and socioemotional functioning in young Chinese children: The moderating role of receptive vocabulary. Early Education and Development, 30(5), 590-607.

Appendices

Appendix A: Residuals versus fitted plot



Appendix B: Q-Q plot of residuals