



Published in final edited form as:

Am J Intellect Dev Disabil. 2012 March ; 117(2): 134–155. doi:10.1352/1944-7558-117.2.134.

Longitudinal Assessment of Intellectual Abilities of Children with Williams Syndrome: Multilevel Modeling of Performance on the Kaufman Brief Intelligence Test-2

Carolyn B. Mervis,

Department of Psychological and Brain Sciences, University of Louisville

Doris J. Kistler,

Department of Psychological and Brain Sciences, University of Louisville

Angela E. John, and

Department of Psychological and Brain Sciences, University of Louisville

the MIND Institute, University of California, Davis.

Colleen A. Morris

Department of Pediatrics, University of Nevada School of Medicine.

Abstract

Multilevel modeling was used to address the longitudinal stability of standard scores (SSs) measuring intellectual ability for children with Williams syndrome (WS). Participants were 40 children with genetically-confirmed WS who completed the Kaufman Brief Intelligence Test-2 (KBIT-2; Kaufman & Kaufman, 2004) 4–7 times over a mean of 5.06 years. Mean age at first assessment was 7.44 years (range: 4.00–13.97 years). On average KBIT-2 Composite IQ, Verbal SS, and Nonverbal SS were stable from 4–17 years, although there were significant individual differences in intercept (Composite IQ, Verbal SS, Nonverbal SS) and slope (Composite IQ, Nonverbal SS). Maternal education was significantly related to Verbal SS intercept. No significant sex differences were obtained. Implications for studies of genotype/phenotype correlations in WS are discussed.

Keywords

Williams syndrome; intellectual disability; intelligence; longitudinal

Williams syndrome (WS) is a rare genetic disorder caused by a microdeletion of 26 genes on chromosome 7q11.23 (Hillier et al., 2003). The prevalence of this syndrome has been estimated at 1 in 7500 live births (Strømme, Bjørnstad, & Ramstad, 2002). WS is associated with specific physical and medical characteristics including a particular facial appearance, cardiovascular disease (especially supravalvar aortic stenosis), connective tissue abnormalities, failure to thrive or growth deficiency, and infantile hypercalcemia (Morris, 2006). Young children with WS have developmental delay; full-scale IQs for most older children and adults are in the borderline to moderate intellectual disability range, although scores range from average for the general population to severe intellectual disability (Mervis & John, 2010). Findings reported in the seminal study of children and adolescents with WS (Bellugi et al., 1988) suggested that despite significant intellectual disability, the language

abilities of individuals with WS were at or close to the level expected for their chronological age (CA). Reports of these findings by secondary sources went further, arguing that the language abilities of children with WS were at least in the average range (e.g., Jackendoff, 1994; Piattelli-Palmarini, 2001). Results from large-sample studies contradict this position. For example, Mervis and John (2010) reported that for a sample of 129 4 – 17-year-olds with WS, mean standard score (SS) was 81.84 on the Peabody Picture Vocabulary Test-4 (PPVT-4; Dunn & Dunn, 2007) and 79.43 on the Expressive Vocabulary Test-2 (EVT-2; Williams, 2007). Both of these SSs are in the test authors' "moderately low" range, indicating that vocabulary is typically not at the level expected for CA. At the same time, Mervis and John (2010) found that mean performance of a sample of 120 4 – 17-year-olds on the Differential Ability Scales, second edition (DASII; Elliott, 2007) was in the borderline range of intellectual ability for both Verbal cluster SS (74.06) and Nonverbal Reasoning cluster SS (78.89); significant intellectual disability was found only on the Spatial cluster (mean SS: 54.86).

Researchers studying individuals with WS in the US typically use the Kaufman Brief Intelligence Test, second edition (KBIT-2; Kaufman & Kaufman, 2004) or its predecessor, the Kaufman Brief Intelligence Test (KBIT; Kaufman & Kaufman, 1990), to describe participants' intellectual abilities (e.g., Key & Dykens, 2010; Lakusta, Dessalegn, & Landau, 2010; Palomares, Englund, & Ahlers, 2011; Plesa Skwerer, Verbalis, Schofield, Faja, & Tager-Flusberg, 2006). The KBIT and the KBIT-2 assess verbal abilities and nonverbal reasoning abilities but do not assess visuospatial construction. The largest sample of individuals with WS for whom performance on the KBIT-2 has been reported (Mervis & Morris, 2007) was composed of 99 4 – 16-year-olds. Mean composite IQ was 72.34 (*SD*: 13.72, range: 40 [floor] – 105.). Consistent with Mervis and John's (2010) findings for the DAS-II, mean Verbal SS was 74.99 (*SD*: 13.92, range: 40 – 108) and mean Nonverbal SS was 76.52 (*SD*: 13.47, range: 47 – 105).

Descriptive statistics such as these are reported for KBIT-2 performance in many studies of individuals with WS, albeit for considerably smaller sample sizes even though the age range is often even broader. Provision of such statistics, and the common practice of treating participants with WS across a very wide age range as a single group, suggests that the performance of individuals with WS on the KBIT-2 may be expected to be consistent across a broad age range. That is, there is not a meaningful increase or decrease in IQ for individuals with WS across the specified age range. Instead, throughout this age range, a person with WS may be expected to continue to develop at a consistent rate relative to his or her general-population age-peers. This assumption has not been addressed empirically. To determine if the assumption is correct, longitudinal studies are critical. In the present study, we used multilevel modeling (Raudenbush & Bryk, 2002; Singer & Willet, 2003) to provide the first longitudinal examination of the consistency of KBIT-2 Composite IQ, Verbal SS, and Nonverbal SS over time for children with WS.

Multilevel models provide both estimates of individual change over time and estimates of inter-individual variability in change trajectories. Additionally, these techniques have the capacity to handle unbalanced designs in which the age at the initial measurement and the measurement intervals may vary. The data described here constitute an unbalanced design since participants entered the study at different CAs and were tested at different intervals. In the remainder of the introduction, we briefly consider the findings of the few previous longitudinal studies of the intellectual abilities of children and adults with WS. As we will consider the possibility of an effect of maternal education level on children's performance on the KBIT-2, we also briefly discuss prior literature addressing the effect of this variable on child IQ for children who do not have WS. (No data addressing the possible role of maternal

education level in child IQ have been reported for children with WS.) We then provide a short introduction to the present study.

Longitudinal Studies of Intellectual Abilities of Individuals with WS

Five longitudinal studies of the intellectual abilities of individuals with WS have been published. In each study, participants' intellectual abilities were measured at two time points and the resulting IQs compared. These studies are summarized in Table 1. As indicated in the table, three studies addressed development in childhood, with mixed findings. In the first study (Crisco, 1990), mean IQ was almost identical at both time points. In a later study (Fisch et al., 2010), which used a more recent version of the same assessment, the mean IQ at the two time points differed by 2 points. The authors reported that IQs were stable for girls but declined significantly for boys. In the third study (Gosch & Pankau, 1996), IQs were stable on one measure (Draw a Person Test; Zeiler, 1971) but declined significantly on the other measure (Columbia Mental Maturity Scale [CMMS]; Bondy, Cohen, Eggert, & Luer, 1969). The CMMS is not normed for the full age range included at Time 2; it is unclear how IQs were computed for the children who were outside the age range for the standardization sample.

In the two other longitudinal studies, participants were adults at least at Time 2 and different assessments were used at the two time points. In the first study (Udwin, Davies, & Howlin, 1996), IQ scores increased significantly from Time 1 to Time 2, especially for the individuals who earned the lowest IQs at Time 1. As the authors point out, these apparent increases are likely spurious; the results of several studies comparing the performance of individuals with borderline intellectual ability or intellectual disability on the WISC-R (Wechsler, 1976) to their performance on the WAIS-R (Wechsler, 1981) have indicated significant increases in WAIS-R IQs relative to WISC-R IQs (see discussion in Searcy et al., 2004), with the amount of increase strongly negatively correlated with WISC-R IQ (e.g., Spitz, 1988). In the second longitudinal study of adults (Howlin, Udwin, Elison, & Stinton, 2010), WAIS-III IQs at Time 2 were lower than WAIS-R IQs at Time 1. However, when the Time 2 IQs were adjusted to account for the fact that for the norming sample, mean scores on the WAIS-III were slightly lower than those for the WAIS-R—by 2.9 points for full-scale IQ, 1.2 points for Verbal IQ, and 4.8 points for Performance IQ (Wechsler, 1997)—mean full-scale IQ and mean Verbal IQ did not differ significantly from Time 1 to Time 2 and mean Performance IQ increased slightly but significantly. The authors noted that the increase was within the 95% confidence interval.

In summary, the results of previous longitudinal studies of IQ in individuals with WS have been mixed, with some reporting significant increases over time, others reporting significant declines, and still others reporting no change. In general, the studies in which significant changes in IQ were reported had clear methodological problems; either the IQ measure at Time 1 was different from the IQ measure at Time 2 or the measure was not normed for the full age range included in the study. Significant changes in IQ were reported in only one study for which the same measure was used at both time points (Fisch et al., 2010). Fisch et al. also reported a significant gender difference, with a decline in IQ for boys but not girls, but did not provide separate IQ means for boys and girls with WS at either time point. In addition, sample sizes in the three studies of children were small, ranging from 14 – 18. In all five studies, only two data points were reported for each participant, so a fine-grained view of any changes in IQ over the period studied was not possible.

Relation between Maternal Educational Level and Child IQ

The relation of maternal education level to child IQ has been considered in a number of studies (see Table 2), some involving children in the general population and others involving

children at risk due to low birth weight or preterm birth. In most studies, maternal education was treated as a dichotomous variable. The definition of “high” maternal education varied, ranging from at least a high school diploma (e.g., Ment et al., 2003) to at least 3 years of college (Leversen et al., 2011). As indicated in Table 2, most studies focused on full-scale IQ and most of these found a significant effect of maternal education, with mean IQ higher for the children in the maternal high education group than in the maternal low education group. The studies that included Verbal IQ and Performance IQ in addition to full-scale IQ found significant effects for all three variables, with the largest effect for Verbal IQ. The few studies that considered only Verbal IQ all found a significant effect for level of maternal education. Finally, in the study that considered Verbal IQ and Performance IQ but did not consider full scale IQ, a significant effect of level of maternal education (treated as a continuous variable) was found for Verbal IQ but not for Performance IQ. In summary, the findings suggest that the effect of maternal education on intellectual abilities is strongest for Verbal IQ but is often also found for full-scale IQ (especially in studies with very large sample sizes) and sometimes found for Performance IQ.

The Present Study

In the present study, we focused on a group of 40 children with WS, all with genetically-confirmed classic deletions, who had completed the KBIT-2 at least 4 times over a period of 3 – 7 years. The relatively large number of data points per child and the broad range of ages at entry into the study allowed us to use multilevel modeling to address the nature of the growth curves for KBIT-2 Composite IQ, Verbal SS, and Nonverbal SS over the period from age 4 – 17 years. We also were able to address the possibility of sex differences and effects of maternal education level on these growth curves.

Method

Participants

Participants were 40 children (21 girls, 19 boys) with genetically confirmed WS who had completed the KBIT-2 between 4 and 7 times (mean: 5.55 times, SD: 1.13, median: 6.0). At least 11.75 months elapsed between administrations. All children had classic-length deletions. At the start of the study children ranged in age from 4.00 – 13.97 years (mean: 7.44 years, SD: 3.17, median: 7.0). Children have participated in the study for 3.00 – 6.25 years (mean: 5.06 years, SD: 1.0, median: 5.13) and ranged in age from 7.07 – 19.09 years (mean: 12.50 years, SD: 3.62, median: 11.99) at their most recent assessment. The average of the children's mean ages during the study (the “grand mean”) was 9.93 years (SD: 3.30, range: 5.56 – 16.48). In the analyses below, the grand mean was rounded to 10.0 years. The racial/ethnic composition of the sample was: 87.5% white non-Hispanic, 5.0% white Hispanic, 2.5% African American non-Hispanic, 2.5% biracial (Asian-white) non-Hispanic, and 2.5% biracial (Asian-white) Hispanic. Thirty of the participants were included in the cross-sectional KBIT-2 sample reported in Mervis and Morris (2007). Twelve mothers did not have college degrees while 13 mothers had bachelor degrees and 15 mothers had advanced degrees.

Materials and Procedure

The Kaufman Brief Intelligence Test, second edition (KBIT-2, Kaufman & Kaufman, 2004) is an individually-administered assessment of intellectual ability normed for ages 4 – 89 years. Both verbal ability and nonverbal reasoning ability are assessed. The Verbal scale includes two subtests: Verbal Knowledge, which measures receptive vocabulary and general knowledge about the world; and Riddles, which measures primarily expressive reasoning. The Nonverbal scale (Matrices) measures nonverbal matrix reasoning, that is, understanding

of relations between either concrete stimuli (pictures of objects) or abstract stimuli (e.g., designs or symbols). The KBIT-2 yields a Verbal SS, a Nonverbal SS, and a Composite IQ based on performance on the Verbal and Nonverbal scales. Each of these has a general-population mean of 100, SD of 15, and range from 40 – 160. According to the KBIT-2 manual (Kaufman & Kaufman, 2004), internal consistency reliability for children aged 4 – 18 years was .90 for the Verbal scale, .86 for the Nonverbal scale, and .92 for IQ Composite. Adjusted test-retest reliability for ages 4 – 12 years was .88 for the Verbal scale, .76 for the Nonverbal scale, and .88 for IQ Composite. Reliability was slightly higher for ages 13 – 21 years. No significant sex differences were reported for Verbal SS or IQ Composite. For Nonverbal SS a small but significant difference in favor of girls was found for ages 7 – 12 years only; the differences for ages 4 – 6 and 13 – 18 years were not significant. The correlation between Verbal SS and Nonverbal SS was .49 for ages 4 – 6 years, .48 for ages 7 – 12 years, and .53 for ages 13 – 18 years. Correlations between performance on the KBIT-2 and performance on the WISC-IV (Wechsler, 2003) for children aged 6 – 16 years were .79 for KBIT-2 Verbal SS and WISC-IV Verbal Comprehension Index, .56 for KBIT-2 Nonverbal SS and WISC-IV Perceptual Reasoning Index, and .77 for KBIT-2 IQ Composite and WISC-IV Full Scale IQ.

The KBIT-2 was administered and scored according to the standardized procedures.

Data Analysis

Multilevel modeling techniques have been used extensively to examine individual differences in longitudinal change (Gibbons, Hedeker, & DuToit, 2010; Raudenbush, 2001; Raudenbush & Bryk, 2002; Singer & Willet, 2003). As the name implies, multilevel models characterize growth (or decline) by specifying models at different hierarchical levels. For a 2-level model, the *Level 1* model describes individual or within-person change over time. The *Level 2* model describes inter-individual variability in the change trajectories defined by the parameters of the *Level 1* model. In the simplest *Level 1* model, individual change over time would be characterized by a linear model which includes only an intercept, typically an indicator of initial status or status at a fixed age, and a slope or change parameter. The *Level 2* model is formulated to account for individual differences in the intercepts and slopes estimated for each individual at *Level 1*. In the present study two predictors (child sex and maternal education level) were introduced at *Level 2* to account for individual variability in IQ intercepts and slopes.

One of the first steps in multilevel modeling is to determine the shape of the function defining the change over time. In our sample some young children showed decline over time while others showed little change or improved. Inspection of individual data suggested that linear change best described the data for the age ranges included in the present sample. To validate this assumption, we tested models that included a quadratic component to model nonlinearity and found no improvement in model fits as indicated by deviance tests. (However, we anticipate that the performance of individuals with steep slopes would stabilize eventually, resulting in nonlinear trends over the life span.) Consequently, the data were modeled by relating KBIT-2 SS to CA centered at the grand sample mean of 10 years. Centering age is recommended because it enables a more meaningful interpretation of the intercept and often increases the stability of the model estimation¹ (Raudenbush & Bryk, 2002; Singer & Willet, 2003). The general-population average SS of 100 was subtracted from each participant's SS so that the model intercepts would provide a statistical comparison of average SS for children with WS to the general-population average.

The *Level 1* model, which estimates an intercept (predicted SS at age 10 years) and slope (rate of annual change in SS) for each participant, is specified in the equation below:

$$SS_{ti} = \pi_{0i} + \pi_{1i}(CA_{ti} - 10) + \varepsilon_{ti} \quad (1)$$

In the *Level 1* equation, SS_{ti} represents the predicted SS – 100 for person i at age t , where age (CA_{ti}) is transformed by subtracting the grand mean age of 10 years. The parameters π_{0i} and π_{1i} are each child's intercept and linear slope, and ε_{ti} is the residual error term indicating the amount of error in predicting the child's SS at a given CA.

Multilevel modeling generally involves testing a number of *Level 2* models. In the simplest *Level 2* model, the individual intercepts and slopes from the *Level 1* model are the outcome variables and the model is formulated to determine if individuals differ in their age 10 estimates (intercepts) and their rates of change (slopes). The *Level 2* model consists of two equations, one for intercept and one for slope:

$$\pi_{0i} = \beta_{00} + r_{0i} \quad (2a)$$

$$\pi_{1i} = \beta_{10} + r_{1i} \quad (2b)$$

In these equations for predicting individual outcomes, β_{00} and β_{10} are estimates of the averages of individual parameter estimates. The parameter averages are also referred to as fixed effects and each can be evaluated by a significance test. The r_{0i} and r_{1i} are person-specific intercept and slope residuals and are assumed to be normally distributed about their respective averages with variances, σ_0^2 and σ_1^2 . These residual variances represent the random effects of the model. These variances are estimates of between-person variability or individual differences in the *Level 1* parameters. Significance tests are available to determine if these variances deviate significantly from a variance of zero. The covariance of intercepts and slopes is also estimated. Intercepts and slopes are often correlated in growth models.

The *Level 2* model described above is often referred to as an “unconditional growth model” because there are no *Level 2* predictors of individual differences in intercepts and slopes. In the present study, we evaluated sex and maternal education as potential *Level 2* predictors of individual intercepts and slopes. As an example, a *Level 2* intercept equation with sex as a predictor is:

$$\pi_{0i} = \beta_{00} + \beta_{01} \text{Sex} + r_{0i} \quad (3)$$

In this model, β_{00} is the average intercept for the males and β_{01} is the predicted average increment (or decrement) relative to the β_{00} estimate for females. A similar equation can be specified for predicting individual slopes, π_{1i} , from sex. The effects of sex and maternal education were evaluated in separate analyses.

The software package HLM 6.08 (Raudenbush, Bryk, Cheong, & Congdon, 2004) was used to fit the multilevel models. We employed full maximum likelihood estimation to enable comparisons of model fit statistics of sex and maternal education models with the unconditional model (i.e., no *Level 2* predictors). Final models were evaluated with

¹It should be noted that centering on the sample mean of 10 years could result in a misspecification of the *Level 1* model due to the fact that the age of 10.0 years is outside the testing range of some of the participants. An alternative is to center the *Level 1* model relative to each person's initial or average age and use the initial or average age as a *Level 2* predictor of the *Level 1* model parameters (Chapman, Hesketh, & Kistler, 2002; Morrell, Brant & Ferrucci, 2009). Although this alternative model might be expected to provide a better fit, it is more complex and estimating a single average developmental trajectory over a broad age range is less straightforward. Using the methods described by Miyazaki and Raudenbush (2000), we compared models using the two methods of centering and obtained comparable fits for each of our outcome variables. Consequently the grand-mean centering model, which is both less complex and more general, was retained.

restricted maximum likelihood (REML) estimation because it is generally thought that REML estimation results in more reliable estimates of variance components when small sample sizes are small (Raudenbush & Bryk, 2002). Although our sample size of 40 is small for a multilevel analysis, simulation studies have shown that fixed effects parameter estimates and standard errors were estimated without bias in sample sizes as small as 30 (e.g., Maas & Hox, 2005). Earlier work by Busing (1993) had suggested that a sample size of 100 was necessary to accurately estimate the *Level 2* variance components. However, Maas and Hox also found that both *Level 1* and *2* variances were estimated reliably although their standard errors were underestimated.

Results

Composite IQ

Parameter estimates and significance tests for the unconditional Composite IQ model are shown in Table 3. The unconditional model indicated a statistically significant average intercept (predicted IQ at the grand mean age of 10 years) but the average slope was not significant. As explained above we subtracted 100 from each SS so that the estimated intercept parameter would reflect the deviation from the general-population Composite IQ of 100. Thus the average intercept of -19 corresponds to a predicted Composite IQ of 81 for a child at the grand mean age of 10 years and is significantly lower than the general-population mean Composite IQ of 100². The estimated average annual change of 0.37 IQ points although in a positive direction, was not statistically significant.

As is often the case in longitudinal studies, there was significant variability in the individual intercepts and slopes. This is indicated by the tests of the random effects or variances, which are also shown in Table 3. The *SDs* of the individual intercepts and slopes (i.e., the residuals r_{0j} and r_{1j}) were 12.8 and 1.0, respectively. Individual intercepts ranged from -51.4 to +16.0. Individual slopes ranged from -1.1 to +2.3. The correlation between Composite IQ intercepts and slopes was 0.42. It should be noted that the random effects variance/covariance estimates may not be highly reliable as population estimates given the relatively small sample size.

The effect of sex as an explanatory variable was evaluated in a second model to determine if there were significant average (i.e., “fixed effect”) differences between males and females in addition to examining individual differences in developmental trajectories as a function of sex. The sex difference in average Composite IQ at age 10 was 0.20 ($p = 0.96$) and the difference in average slope was .26 ($p = 0.63$).

The effects of maternal education were also evaluated. Although not statistically significant ($p = .09$), the estimated age 10 Composite IQ was almost 8 points higher for children whose mothers had 4-year college degrees. Individual differences in slopes were not related to maternal education. Table 4 shows the results of the maternal education model for Composite IQ.

Because the addition of maternal education and sex did not result in significant improvements in the Composite IQ model, the unconditional model was used to estimate the average developmental trajectory and individual trajectories. Figure 1 shows the predicted average trajectory from ages 4 – 17 years and fitted functions for each of the 40 individuals,

²Only one participant ever earned a Composite IQ of 40, the lowest possible IQ on the KBIT-2. This child earned an IQ of 40 on two of her seven assessments. On one occasion, her Verbal SS and Nonverbal SS were both above 40, so she did not perform at floor on either of the scales included in the composite IQ. On the second occasion, her Verbal SS was at floor (40) but her Nonverbal SS was above floor. No other child ever scored at floor on either the Verbal scale or the Nonverbal scale.

estimated over each individual's age span in the study. Although maternal education was not a significant predictor, the data for the two education groups are plotted in different colors to enable comparisons with the results of the Verbal SS and Nonverbal SS models.

Figure 1 emphasizes the individual variability in developmental trajectories. While the average developmental trajectory suggested no significant change in Composite IQ from childhood through adolescence, there were some children who showed substantial decline and some children with notable increases. Figure 2 shows the raw data and fitted functions for 12 quasi randomly selected participants. Individuals were assigned to quartiles based on their estimated slopes. Three individuals who had completed the KBIT-2 at least 5 times were randomly selected from each quartile. As is apparent in the figure, there is considerable variability in the slopes of individual functions and in the variability about the individual function (i.e., lack of fit).

Verbal SS

The parameter estimates and significance tests for the unconditional model for Verbal SS are shown in Table 5. For Verbal SS the estimated average intercepts for a child at the grand mean age of 10 years was -18.5 (i.e., 81.5), which was significantly lower than the general-population average of 100. The average slope of -0.19, indicating a very small annual decrease, was not significantly different from a slope of 0. There was substantial variability in individual developmental trajectories for Verbal SS, although differences in slope variances were not statistically significant. The *SDs* for individual intercepts and slopes were 11.3 and 0.37, respectively. Individual intercepts ranged from -49.7 to +4.3. Individual slopes ranged from -1.3 to +0.5. Verbal SS intercepts and slopes were highly correlated ($r = 0.98$), suggesting that children with higher SSs age 10 showed greater positive change while children with lower SSs showed greater decline.

The fit statistics for the sex model did not indicate a significant improvement. On average, females scored 3.2 points lower ($p = 0.40$) than males at age 10 years but had slightly steeper slopes (0.16, $p = 0.66$). As shown in Table 6, maternal education was a significant predictor of intercept variance but not slope variance. Children of mothers with 4-year college degrees had a predicted age 10 Verbal SS that was 8 points higher (mean = 83.9) than those children whose mothers did not have 4-year college degrees (mean = 75.8). The average annual decline was -0.17 for children of mothers with 4-year college degrees and -0.23 for children of mothers without 4-year college degrees. The average developmental trajectories for Verbal SS for each group are shown in Figure 3. Individual function fits are also shown. Data and model fits for the 12 individuals included in Figure 2 are shown in Figure 4 for Verbal SS. The inclusion of maternal education as a factor resulted in an estimated 8% $[(127.7-117.13)/127.71]$ reduction in intercept variance when compared to the unconditional model (Raudenbush & Bryk, 2002).

Nonverbal SS

The statistics for the unconditional model for Nonverbal SS are shown in Table 7. The average intercept was -13.9, corresponding to a predicted Nonverbal SS of 86.1 at the grand mean age of 10 years, is significantly lower than the general-population average SS of 100. The average slope of 0.72 was not significantly different from a slope of 0 indicating no significant average change in Nonverbal SS. The intercept and slope standard deviations were 13.0 and 1.9, respectively. Individual intercepts ranged from -41.6 to +21.9. Individual slopes ranged from -2.9 to +4.9. Tests of intercept and slope variability indicated significant individual differences in developmental trajectories.

Sex was not a significant predictor of intercept and slope variability. The estimated average Nonverbal SS for a female at age 10 was 3.9 points higher ($p = 0.39$) than the male average. The average slope was 0.44 points higher ($p = 0.59$) for females than for males. As shown in Table 8, maternal education also was not a significant predictor of Nonverbal SS developmental trajectories. Predicted scores at age 10 were approximately 4.6 points higher and slopes were 0.17 points greater for children of college-educated mothers. The average developmental trajectory for Nonverbal SS, shown in Figure 5, was based on the unconditional model. Individual curves are plotted in different colors for the two maternal education groups to facilitate comparisons with the Verbal SS results. Individual model fits for Nonverbal SS are shown in Figure 6 for the 12 participants included in Figures 2 and 4.

Relations between Verbal SS and Nonverbal SS trajectories

To examine the relations between the developmental trajectories for Verbal SS and Nonverbal SS for individual children, we computed two Spearman correlations, one between the estimated individual intercepts for the Verbal and Nonverbal scales and the other between the estimated individual slopes for these scales. Spearman's rho (r_s) was used because several outliers were present. As shown in Figure 7, the Verbal SS and Nonverbal SS intercepts, which estimate the child's SSs at age 10 years, were moderately correlated ($r_s = 0.51$, $p < 0.008$). As shown in Figure 8, the correlation between Verbal and Nonverbal slope estimates was considerably smaller ($r_s = 0.12$, $p = 0.033$). The variability in slopes for the Nonverbal SS was considerably greater than the Verbal SS slope variances as is apparent in Figure 8.

Discussion

The growth curve model findings from this study indicate that on average KBIT-2 Composite IQ, Verbal SS, and Nonverbal SS are stable across the age range from 4 – 17 years for children with WS. At the same time, there were significant individual differences among children with WS for predicted SS at age 10 years (intercept) for all three measures. In addition, there were significant individual differences for annual rate of change in SS (slope) for both Composite IQ and Nonverbal SS. In contrast to earlier suggestions that the verbal abilities of individuals with WS are in the average range for the general population and are stronger than their nonverbal abilities, mean intercept for Verbal SS was 81.5 (considerably below the general population mean of 100) and was 4.6 points lower than mean intercept for Nonverbal SS (86.1). This pattern is consistent with that found by Mervis and John (2010) for performance on the DAS-II Verbal and Nonverbal Reasoning clusters and provides further support for their argument that children with WS have specific difficulty with the spatial domain rather than with nonverbal ability in general. In the remainder of the Discussion we first consider possible bases for the individual differences in intercepts and slopes that we obtained. We then relate our findings to prior findings from longitudinal studies both of individuals with WS and individuals with other genetic syndromes. Next, we consider the implications of our results for the use of the KBIT-2 in studies of genotype/phenotype correlations in Williams syndrome. Finally, we consider some of the limitations of this study.

Possible Bases for Significant Individual Differences in KBIT-2 Intercepts or Slopes

The significant individual differences that were found in the present study were not due to differences between children in the number of genes deleted or the specific genes deleted. All participants had classic deletions so all had the same 26 genes on chromosome 7q11.23 deleted. It was still possible, however, that the lowest performing children had genetic “second hits.” That is, in addition to the deletion causing WS, these children might have had one or more additional deletions or duplications that resulted in a further negative impact on

intellectual ability. To consider this possibility, Affymetrix Genome-Wide Human SNP Array 6.0 microarray analyses and follow-up fluorescence *in situ* hybridization (FISH) tests were performed for nine participants including the two who earned the lowest SSs, both of whom scored in the range of moderate intellectual disability. No additional deletions or duplications were confirmed. Thus, it is unlikely that the individual differences in intellectual ability that we found were due to deletions or duplications outside of the WS region.

Two other possibilities were addressed in the models reported in the Results. The first was that girls performed significantly differently from boys. However, the model that included Sex as a predictor did not result in a significant improvement in fit over the unconditional model for Composite IQ, Verbal SS, or Nonverbal SS, indicating that the significant individual differences were not due to differences between the performance of girls and the performance of boys.

The second possibility we considered was that individual differences in SS intercept or slope were related to maternal education level (or factors associated with this variable). Consistent with prior findings for studies of children in the general population or other special populations, children with WS whose mothers had a higher level of education (operationalized in this study as having a bachelor or more advanced degree) had a significantly higher Verbal SS intercept. Also consistent with prior findings, there was no effect of maternal education on slope (rate of SS change) for Composite IQ, Verbal IQ, or Nonverbal IQ. Maternal education level was not related to child Nonverbal SS intercept. Prior findings are mixed on this point, but in studies that did find a significant effect of maternal education level on child Performance IQ (which includes both nonverbal reasoning and visuospatial construction), the effect was consistently smaller than for Verbal IQ. In contrast to prior findings, in the present study maternal education level also was not significantly related to child Composite IQ. The sample sizes in prior studies, however, were considerably larger than in the present study. It is possible that with a larger sample size, we would have found a significant effect of maternal education level on child Composite IQ. Note, however, that in our study as in previous studies, such a finding would have been driven primarily by the effect of maternal education level on Verbal IQ.

Comparisons with Prior Longitudinal Studies

The present study differed from prior longitudinal studies of the intellectual abilities of individuals with WS in that children were tested at least four times with the same assessment whereas in previous studies participants were tested with the same assessment at most twice. The results of prior studies were mixed, although most indicated that on average, SSs for individuals with WS were stable across the age range tested, particularly after adjustments for the use of different editions of an assessment were made. Only one of the five studies reported a different pattern of results as a function of sex, with boys showing a significant decline in IQ while girls were stable. Thus, the findings of prior studies are broadly consistent with those of the present study, which included a sample size approximately twice that of previous longitudinal studies of children with WS and collected more than twice as many data points per participant.

Many longitudinal studies of the intellectual abilities of children who have other genetic syndromes associated with intellectual disability have been published. However, we were able to find only six studies that used SSs rather than age-equivalent (AE) scores as the dependent variable and that administered the same assessment at all time points. (For a discussion of problems with the use of AE scores see Mervis & Klein-Tasman, 2004 and Mervis & Robinson, 2005.) The participants in three of the studies had genetically-confirmed fragile X syndrome. Hagerman et al. (1989) compared the first and last Stanford-

Binet L-M IQs for 24 males with fragile X syndrome (20 of whom were with the age range for which the assessment was normed). Mean CA was 11.42 years at the first assessment and 16.27 years at the end of the study. At the group level, a significant decline in IQ was found. However, only 7 of the 20 individuals who were within the age-range of the norms evidenced a significant decline. Fisch et al. (2010) compared the Stanford-Binet IV IQs of 25 boys and 12 girls with fragile X syndrome at two time points an average of 2 years apart. A significant decline in IQ was found for boys but not for girls. Skinner et al. (2005) used hierarchical linear modeling to consider longitudinal changes in the Leiter International Performance Scale-Revised (Roid & Miller, 1997) Brief IQs of 45 boys with fragile X syndrome tested annually a mean of 3.5 times (range: 2 – 6). At the start of the study children's ages ranged from 4.0 – 13.8 years. A significant linear decline was found, averaging 4.2 points per year, with a significant reduction in slope after age 8 years. Maternal education level was not related to Leiter-R Brief IQ intercept or slope.

The participants in the remaining studies had either Lesch-Nyhan syndrome or Costello syndrome. Matthews, Solan, Barabas, and Robey (1999) compared the Stanford-Binet IV IQs for 6 boys with genetically-confirmed Lesch-Nyhan syndrome, another X-linked genetic disorder, at three time points across a 4-year period. Mean age at the end of the study was 17.83 years. Mean IQ at Time 3 was significantly lower than at either Time 1 or Time 2. Axelrad, Nicholson, Stabley, Sol-Church, and Gripp (2007) compared the Leiter-R Brief IQs for 12 children and adolescents with genetically-confirmed Costello syndrome at two time points two years apart. Participants' ages ranged from 2.1 – 17.8 years at Time 1. The difference between the two sets of IQs was not significant. Axelrad et al. (2009) added a third time point (2 years after the second) for 11 of the participants and reported no significant difference between any pair of time points for Leiter-R Brief IQ. No sex differences were found. In conjunction with our results for children with WS, these findings suggest, not surprisingly, that the mean trajectory for IQ over time may differ both as a function of syndrome and as a function of sex (especially for X-linked syndromes). In particular, similarly to the findings we reported for both girls and boys with WS, mean IQ for both boys and girls with Costello syndrome did not differ significantly over the three time points measured during a 4-year period and mean IQ for girls with fragile X syndrome did not differ significantly at the start and end of the 2-year period studied by Fisch et al. In contrast, mean IQ for boys with either fragile X syndrome or Lesch-Nyhan syndrome declined significantly over time. Similarly to our finding for KBIT-2 Nonverbal SS, maternal educational level was not related to either the intercept or the slope for the Leiter-R Brief IQ, a measure of nonverbal ability, for boys with fragile X syndrome.

Implications for Studies of Genotype/Phenotype Relations

Studies of genotype/phenotype relations almost always require large sample sizes. As a result, at least for rare syndromes, it is not practical to restrict the age range in a single study to a narrow interval or to a single sex. Instead, by necessity, both males and females will be included and the participants are likely to vary considerably in age. If the entire age range will be treated as a single group, then it is important that performance on the variable measuring the phenotypic trait be independent of CA for that syndrome. For example, the deletion that causes WS has a maternal origin in approximately 50% of cases and a paternal origin in the remaining 50% (Hobart et al., 2010). If one wished to address the question of whether the IQ of children with WS differed as a function of the parent-of-origin of the deletion that caused WS, it would likely be necessary to include both males and females across a wide age range in order to have an adequate sample size. In such a case, it would be important to use an IQ measure for which average performance for individuals with WS does not vary as a function of CA or sex. Use of such a measure also would permit the researcher to address the question of whether any effect of parent-of-origin on IQ differs as a

function of the sex of the individual with WS. Many other genotypic variables also could be examined. For example, the researcher could address the possibility that Composite IQ, Verbal SS, and/or Nonverbal SS differ as a function of which allele of one of the single-copy genes the child has or as a function of which genes are deleted. The results of the present study indicate that for children with WS aged 4 – 17 years, KBIT-2 Composite IQ, Verbal SS, and Nonverbal SS do not vary significantly either as a function of age or sex. Thus, these measures are appropriate for use in addressing possible genotype/phenotype relations involving IQ, verbal ability, and/or nonverbal reasoning ability for children with WS.

Limitations

Although the sample size of the present study is approximately twice that of previous longitudinal studies of the intellectual abilities of children with WS, it is still relatively small. In order to address the possible impact of additional factors on individual differences in children's intellectual abilities and in order to unpack the maternal education variable into its component variables (e.g., variety and amount of stimulation, contingent vs. noncontingent delivery of stimulation, socio-economic status, material resources, family stress, supports available to the family; see for example Bee et al., 1982), it will be important to increase the number of children in this or future studies. Similarly, although each child was tested 4 – 7 times with the same measure (at least twice as often as in previous studies) the time period covered for an individual child was relatively short (3 – 6 years) and we only were able to model SSs for the age period from 4 – 17 years. Collection of additional data points, especially at older ages, would allow for modeling over a wider age range and for more accurate specification of the shape of SS growth. Although growth from age 4 – 17 years was best described as linear, changes in SS performance almost certainly stabilize at some point, resulting in a nonlinear trend. Specification of this trend likely will require collection of data from older adolescents and young adults. Finally, the KBIT-2 only includes scales measuring verbal ability and nonverbal reasoning ability. To determine if SSs for visuospatial construction (the area of greatest weakness for individuals with WS) are stable across childhood, longitudinal studies using measures of intellectual ability that specifically address visuospatial construction are needed. Replication of the findings for verbal ability and nonverbal reasoning ability using other assessments also is important. To begin to address these concerns, we currently are conducting longitudinal studies of the SS performance of children with WS on the DAS-II (Elliott, 2007), the PPVT-4 (Dunn & Dunn, 2007), and the EVT-2 (Williams, 2007).

Conclusion

The results of the present longitudinal study indicate that mean KBIT-2 Composite IQ, Verbal SS, and Nonverbal SS all are stable over time for children with WS between 4 and 17 years of age. This finding, combined with the lack of sex differences on any of these measures, indicates that the KBIT-2 has considerable value for studies addressing topics such as genotype/phenotype correlations in WS where it is likely that both boys and girls will be included and that participants will vary considerably in age due to the rarity of the syndrome and the need for samples to be as large as possible. Future studies that include larger numbers of participants over a wider age range so that the SS growth functions may be specified more precisely and additional factors affecting individual differences in KBIT-2 Composite IQ, Verbal SS, and Nonverbal SS identified would further enhance the value of the KBIT-2 for research on individuals with WS.

Acknowledgments

This research was supported by grants R37 HD29957 from the National Institute of Child Health and Human Development and R01 NS35102 from the National Institute of Neurological Disorders and Stroke. We are grateful to the participants and their families for their long-term commitment to our research. We thank the members of our laboratory who administered the Kaufman Brief Intelligence Test-2 to the participants in this study.

References

- Axelrad ME, Nicholson L, Stabley DL, Sol-Church, Gripp KW. Longitudinal assessment of cognitive characteristics in Costello syndrome. *American Journal of Medical Genetics*. 2007; 143A:3185–3193. doi:10.1002/ajmg.a.31968. [PubMed: 17963256]
- Axelrad ME, Schwartz DD, Fehlis JE, Hopkins E, Stabley DL, Gripp KW. Longitudinal course of cognitive, adaptive, and behavioral characteristics in Costello syndrome. *American Journal of Medical Genetics*. 2009; 149A:2666–2672. doi:10.1002/ajmg.a.33126. [PubMed: 19919001]
- Bee HL, Barnard KE, Eyres SJ, Gray CA, Hammond MA, Spietz AL, Clark B. Prediction of IQ and language skill from perinatal status, child performance, family characteristics, and mother-infant interaction. *Child Development*. 1982; 53:1134–1156. doi:10.2307/1129003. [PubMed: 7140423]
- Bellugi, U.; Marks, S.; Bihrlle, A.; Sabo, H. Dissociation between language and cognitive functions in Williams syndrome. In: Bishop, D.; Mogford, K., editors. *Language development in exceptional circumstances*. Churchill Livingstone; London: 1988. p. 177-189.
- Bondy, C.; Cohen, R.; Eggert, D.; Luer, G. *Testbatterie für Geistig Behinderte Kinder TBGB*. Beltz-Verlag; Weinheim, Germany: 1969.
- Breslau N, Chilcoat HD, Susser ES, Matte T, Liang K-Y, Peterson EL. *American Journal of Epidemiology*. 2001; 154:711–717. doi:10.1093/aje/154.8.711. [PubMed: 11590083]
- Busing, FMTA. *Distribution characteristic of variance estimates in two-level models: a Monte-Carlo study*. Leiden University, Department of Psychology; 1993. Technical Report PRM 93-04
- Chapman R, Hesketh LJ, Kistler D. Predicting longitudinal change in language production and comprehension in individuals with Down Syndrome: hierarchical linear modeling. *Journal of Speech, Language, and Hearing Research*. 2002; 45:902–915. doi:10.1044/1092-4388(2002/073).
- Cornelius MD, Goldschmidt L, De Genna NM, Richardson GA, Leech SL, Day R. Improvement in intelligence test scores from 6 to 10 years in children of teenage mothers. *Journal of Developmental and Behavioral Pediatrics*. 2010; 31:405–413. doi:10.1097/DBP.0b013e3181e121d2. [PubMed: 20495472]
- Crisco JJ. Rate of cognitive development in young children with Williams syndrome. *Clinical Research*. 1990; 38:536A.
- Dunn, LE.; Dunn, DM. *Peabody Picture Vocabulary Test*. 4th edition. Pearson Assessments; Minneapolis, MN: 2007.
- Elliott, CD. *Differential Ability Scales*. 2nd edition. Psychological Corporation; San Antonio, TX: 2007.
- Fisch GS, Carpenter N, Howard-Peebles PN, Holden JJA, Tarleton J, Simensen R. The course of cognitive-behavioral development in children with the FMR1 mutation, Williams-Beuren syndrome, and neurofibromatosis type 1: The effect of gender. *American Journal of Medical Genetics Part A*. 2010; 152A:1498–1509. doi:10.1002/ajmg.a.33412. [PubMed: 20503326]
- Gale CR, O'Callaghan FJ, Godfrey KM, Law CM, Martyn CN. Critical periods of brain growth and cognitive function in children. *Brain*. 2004; 127:321–329. doi:10.1093/brain/awh034. [PubMed: 14645144]
- Gibbons RD, Hedeker D, DuToit S. Advances in analysis of longitudinal data. *Annual Review of Clinical Psychology*. 2010; 6:79–107. doi:10.1146/annurev.clinpsy.032408.153550.
- Gosch A, Pankau R. Longitudinal study of the cognitive development in children with Williams-Beuren syndrome. *American Journal of Medical Genetics*. 1996; 61:26–29. doi:10.1002/(SICI)1096-8628(19960102)61:1<26::AID-AJMG5>3.3.CO;2-V. [PubMed: 8741913]
- Hagerman RJ, Schreiner RA, Kemper MB, Wittenberger MD, Zahn B, Habicht K. Longitudinal IQ changes in fragile X males. *American Journal of Medical Genetics*. 1989; 33:513–518. doi:10.1002/ajmg.1320330422. [PubMed: 2596514]

- Hedrick, DL.; Prather, EM.; Tobin, AR. *Sequenced Inventory of Communication Development*. University of Washington Press; Seattle, WA: 1975.
- Hillier LW, Fulton RS, Fulton LA, Graves TA, Pepin KH, Wagner-McPherson C, Wilson RK. The DNA sequence of chromosome 7. *Nature*. 2003; 424:157–164. doi:10.1038/nature01782. [PubMed: 12853948]
- Hobart HH, Morris CA, Mervis CB, Pani AM, Kistler DJ, Rios CM, Bray-Ward P. Inversion of the Williams syndrome region is a common polymorphism found more frequently in parents of children with Williams syndrome. *American Journal of Medical Genetics Part C*. 2010; 154C:220–228. doi:10.1002/ajmg.c.30258.
- Howlin P, Elison S, Udwin O, Stinton C. Cognitive, linguistic and adaptive functioning in Williams syndrome: Trajectories from early to middle adulthood. *Journal of Applied Research in Intellectual Disabilities*. 2010; 23:322–336. doi:10.1111/j.1468-3148.2009.00536.x.
- Jackendoff, R. *Patterns in the mind: Language and human nature*. Basic; New York: 1994.
- Jedrychowski W, Perera F, Jankowski J, Butscher M, Mroz E, Flak E, Sowa A. Effect of exclusive breastfeeding on the development of children's cognitive function in the Krakow prospective birth cohort study. *European Journal of Pediatrics*. in press. DOI 10.1007/s00431-011-1507-5.
- Kaufman, AS.; Kaufman, NL. *Kaufman Brief Intelligence Test*. American Guidance Services; Circle Pines, MN: 1990.
- Kaufman, AS.; Kaufman, NL. *Kaufman Brief Intelligence Test*. second edition. American Guidance Services; Circle Pines, MN: 2004.
- Key APF, Dykens EM. Electrophysiological study of local/global processing in Williams syndrome. *Journal of Neurodevelopmental Disorders*. 2010; 3:28–38. doi: 10.1007/s11689-010-9064-1. [PubMed: 21484595]
- Klebanov P, Brooks-Gunn J. Cumulative, human capital, and psychological risk in the context of early intervention: Links with IQ at ages 3, 5, and 8. *Annals of the New York Academy of Sciences*. 2006; 1094:63–82. doi:10.1196/annals.1376.007. [PubMed: 17347342]
- Lakusta L, Dessalegn B, Landau B. Impaired geometric reorientation caused by genetic defect. *Proceedings of the National Academy of Sciences*. 2010; 107:2813–2817. doi:10.1073/pnas.0909155107.
- Leveresen KT, Sommerfelt K, Rønnestad A, Kaaresen PI, Farstad T, Skranes J, Markestad T. Prediction of neurodevelopmental and sensory outcome at 5 years in Norwegian children born extremely preterm. *Pediatrics*. 2011; 127:e630–638. DOI: 10.1542/peds.2010-1001. [PubMed: 21321031]
- Maas CJM, Hox JJ. Sufficient sample sizes for multilevel modeling. *Methodology*. 2005; 1:86–92. doi: 10.1027/1614-1881.1.3.86.
- Matthews W, Solan A, Barabas G, Robey K. Cognitive functioning in Lesch-Nyhan syndrome: a 4-year follow-up study. *Developmental Medicine & Child Neurology*. 1999; 41:260–262. doi: 10.1017/S0012162299000547. [PubMed: 10355810]
- Ment L, Vohr B, Allan W, Katz KH, Schneider KC, Westerveld M, Makuch RW. Change in cognitive function over time in very low-birth-weight infants. *Journal of the American Medical Association*. 2003; 289:705–711. doi:10.1001/jama.289.6.705. [PubMed: 12585948]
- Mervis CB, John AE. Cognitive and behavioral characteristics of children with Williams syndrome: Implications for intervention approaches. *American Journal of Medical Genetics Part C*. 2010; 154C:229–248. doi:10.1002/ajmg.c.30263.
- Mervis CB, Klein-Tasman BP. Methodological issues in group-matching designs: α levels for control variable comparisons and measurement characteristics of control and target variables. *Journal of Autism and Developmental Disorders*. 2004; 34:7–17. doi:10.1023/B:JADD.0000018069.69562.b8. [PubMed: 15098952]
- Mervis, CB.; Morris, CA. Williams syndrome.. In: Mazzocco, MMM.; Ross, JL., editors. *Neurogenetic developmental disorders: Variation of manifestation in childhood*. MIT Press; Cambridge, MA: 2007. p. 199-262.
- Mervis CB, Robinson BF. Designing measures for profiling and genotype/phenotype studies of individuals with genetic syndromes or developmental language disorders. *Applied Psycholinguistics*. 2005; 26:41–64. doi:10.1017/S0142716405050058.

- Miyazaki Y, Raudenbush SW. Tests for linkage of multiple cohorts in an accelerated longitudinal design. *Psychological Methods*. 2000; 5:44–63. doi:10.1037/1082-989X.5.1.44. [PubMed: 10937322]
- Morrell CH, Brant LJ, Ferrucci L. Model choice can obscure results in longitudinal studies. *Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*. 2009; 64:215–222. doi: 10.1093/gerona/gln024.
- Morris, CA. The dysmorphology, genetics, and natural history of Williams-Beuren syndrome.. In: Morris, CA.; Lenhoff, HM.; Wang, PP., editors. *Williams-Beuren syndrome: Research, evaluation, and treatment*. Johns Hopkins University Press; Baltimore, MD: 2006. p. 3-17.
- Palomares M, Englund JA, Ahlers S. Patterns and trajectories in Williams syndrome: The case of visual object discrimination. *Research in Developmental Disabilities*. 2011; 32:1021–1029. doi: 10.1016/j.ridd.2011.01.038. [PubMed: 21334848]
- Piattelli-Palmarini M. Speaking of learning: how do we acquire our marvellous facility for expressing ourselves in words? *Nature*. 2001; 411:887–888.
- Plesa Skwerer D, Verbalis A, Schofield C, Faja S, Tager-Flusberg H. Social-perceptual abilities in adolescents and adults with Williams syndrome. *Cognitive Neuropsychology*. 2006; 23:338–349. doi:10.1080/02643290542000076. [PubMed: 21049334]
- Raudenbush SW. Comparing personal trajectories and drawing causal inferences from longitudinal data. *Annual Review of Psychology*. 2001; 52:501–525. doi:10.1146/annurev.psych.52.1.501.
- Raudenbush, SW.; Bryk, AS. *Hierarchical linear models: Applications and data analysis methods*. 2nd ed.. Sage; Thousand Oaks, CA: 2002.
- Raudenbush, SW.; Bryk, AS.; Cheong, YF.; Congdon, RT. *HLM 6: Hierarchical linear and nonlinear modeling*. Scientific Software International; Lincolnwood, IL: 2004.
- Roid, GH.; Miller, LJ. *Leiter International Performance Scale-Revised*. Stoelting; Wood Dale, IL: 1997.
- Saavalainen P, Luoma L, Bowler D, Timonen T, Määttä S, Laukkanen E, Herrgård E. Naming skills of children born preterm in comparison with their term peers at the ages of 9 and 16 years. *Developmental Medicine & Child Neurology*. 2006; 48:28–32. doi:10.1017/S0012162206000077. [PubMed: 16359591]
- Searcy YM, Lincoln AJ, Rose FE, Klima ES, Bavar N, Korenberg JR. The relationship between age and IQ in adults with Williams syndrome. *American Journal on Mental Retardation*. 2004; 109:231–236. doi:10.1352/0895-8017(2004)109<231:TRBAAI>2.0.CO;2. [PubMed: 15072519]
- Singer, JD.; Willett, JB. *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford University Press; New York, NY: 2003.
- Skinner M, Hooper S, Hatton DD, Roberts J, Mirrett P, Schaaf J, Bailey DB Jr. Mapping nonverbal IQ in young boys with fragile X syndrome. *American Journal of Medical Genetics*. 2005; 132A:25–32. doi:10.1002/ajmg.a.30353. [PubMed: 15551333]
- Sommerfelt K, Ellertsen B, Markestad T. Parental factors in cognitive outcome of non-handicapped low birthweight infants. *Archives of Disease in Childhood*. 1995; 73:F135–F142. doi:10.1136/fn.73.3.F135. [PubMed: 8535868]
- Spitz HH. Inverse relationship between the WISC-R/WAIS-R score disparity and IQ level in the lower range of intelligence. *American Journal on Mental Retardation*. 1988; 92:376–378. [PubMed: 3342139]
- Strømme P, Bjørnstad PG, Ramstad K. Prevalence estimation of Williams syndrome. *Journal of Child Neurology*. 2002; 17:269–271. doi:10.1177/088307380201700406. [PubMed: 12088082]
- Terman, LM.; Merrill, MA. *Stanford-Binet Intelligence Scale Form L-M*. Riverside Publishing; Rolling Meadows, IL: 1972.
- Thorndike, RL.; Hagen, EP.; Sattler, JM. *The Stanford-Binet Intelligence Scale*. 4th edition. Riverside Publishing; Chicago, IL: 1986.
- Udwin O, Davies M, Howlin P. A longitudinal study of cognitive abilities and educational attainment in Williams syndrome. *Developmental Medicine and Child Neurology*. 1996; 38:1020–1029. doi: 10.1111/j.1469-8749.1996.tb15062.x. [PubMed: 8913183]
- Vohr B, Allan WC, Westerveld M, Schneider KC, Katz KH, Makuch RW, Ment LR. School-age outcomes of very low birth weight infants in the indomethacin intraventricular hemorrhage

prevention trial. *Pediatrics*. 2003; 111:e340–e346. doi:10.1542/peds.111.4.e340. [PubMed: 12671149]

Wechsler, D. Wechsler Preschool and Primary Scale of Intelligence. Psychological Corporation; San Antonio, TX: 1967.

Wechsler, D. Wechsler Intelligence Scale for Children-Revised. NFER-Nelson; Windsor, UK: 1976.

Wechsler, D. Wechsler Adult Intelligence Test-Revised UK edition. Psychological Corporation; London, UK: 1981.

Wechsler, D. Wechsler Intelligence Scale for Children. 3rd edition. Psychological Corporation; San Antonio, TX: 1991.

Wechsler, D. Wechsler Adult Intelligence Test. 3rd edition. Psychological Corporation; San Antonio, TX: 1997.

Wechsler, D. Wechsler Abbreviated Scale of Intelligence. Psychological Corporation; San Antonio, TX: 1999.

Wechsler, D. Wechsler Intelligence Scale for Children. 4th edition. Psychological Corporation; San Antonio, TX: 2003.

Williams, KT. Expressive Vocabulary Test. 2nd edition. Pearson Assessments; Minneapolis, MN: 2007.

Zeiler, H. Der Mann-Zeichen-Test in Detailstatistischer Auswertung. Aschendorffsche Verlagsbuchhandlung; Münster, Germany: 1971.

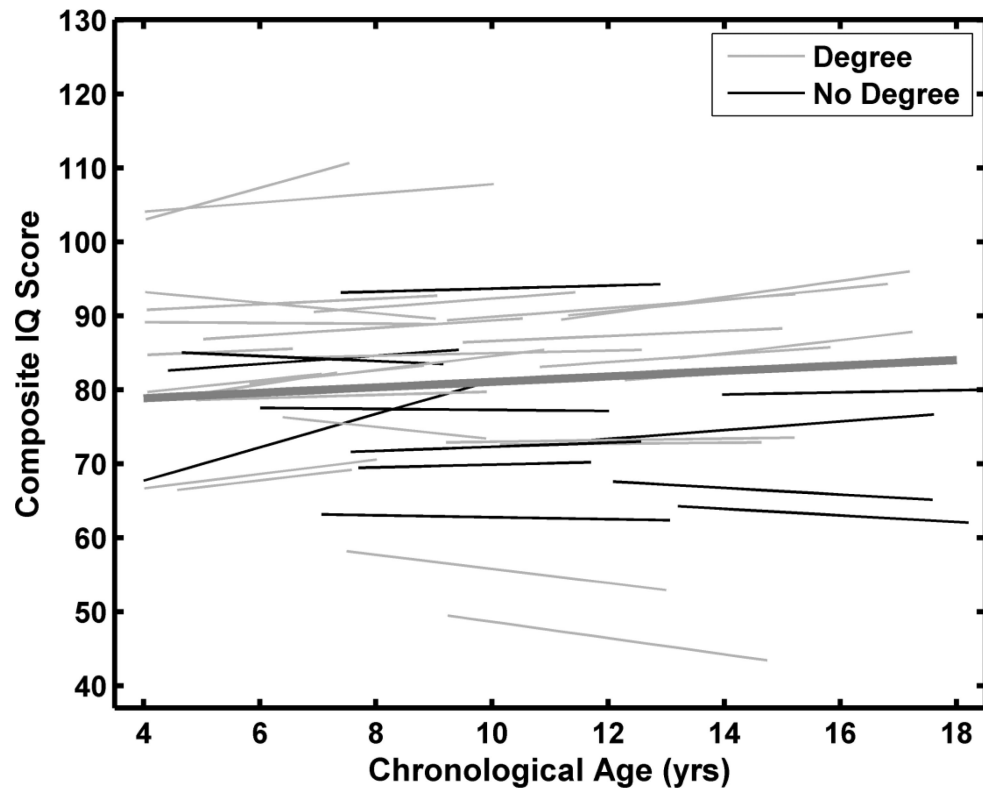


Figure 1. Predicted average developmental trajectory and individual fitted functions for the KBIT-2 Composite IQ scores. The function fits are based on the unconditional model. No degree = High school diploma or some college. Degree = Bachelor or more advanced degree.

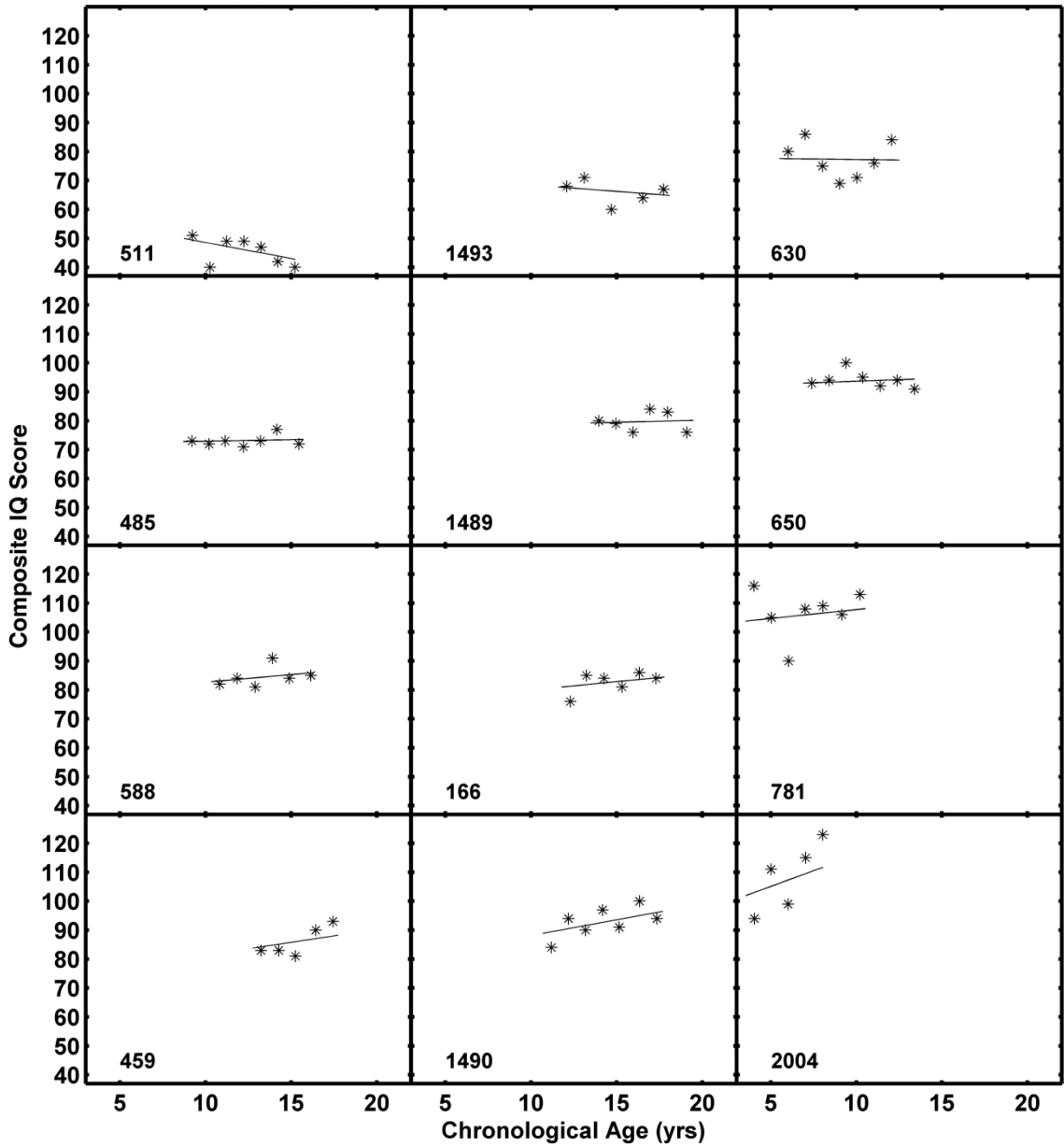


Figure 2. KBIT-2 Composite IQ scores with fitted functions for 12 quasi randomly selected children. Children were ordered by slope and three were randomly selected from each quartile, contingent upon having completed the KBIT-2 at least five times. Each row corresponds to a quartile, with the lowest quartile (i.e., most negative slopes) in the top row and the highest quartile in the bottom row. The numbers in the lower left corner of each panel are participant identifiers.

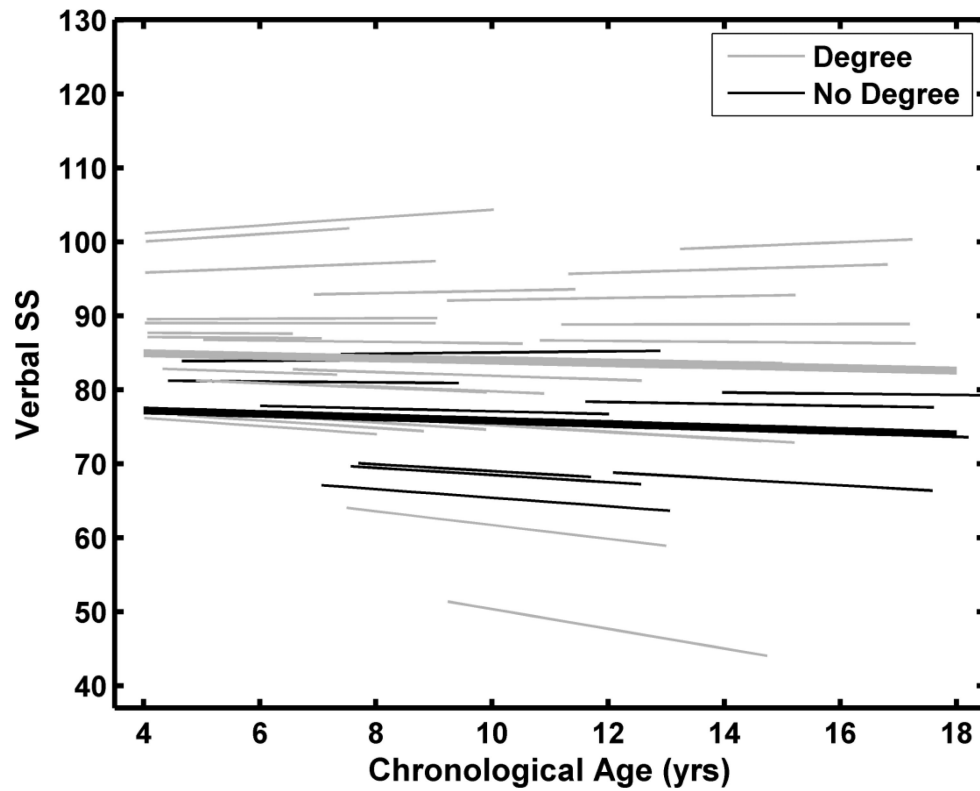


Figure 3. Predicted average developmental trajectories and individual fitted functions for the KBIT-2 Verbal standard scores (SS). The function fits were based on the maternal education model. No degree = High school diploma or some college. Degree = Bachelor or more advanced degree.

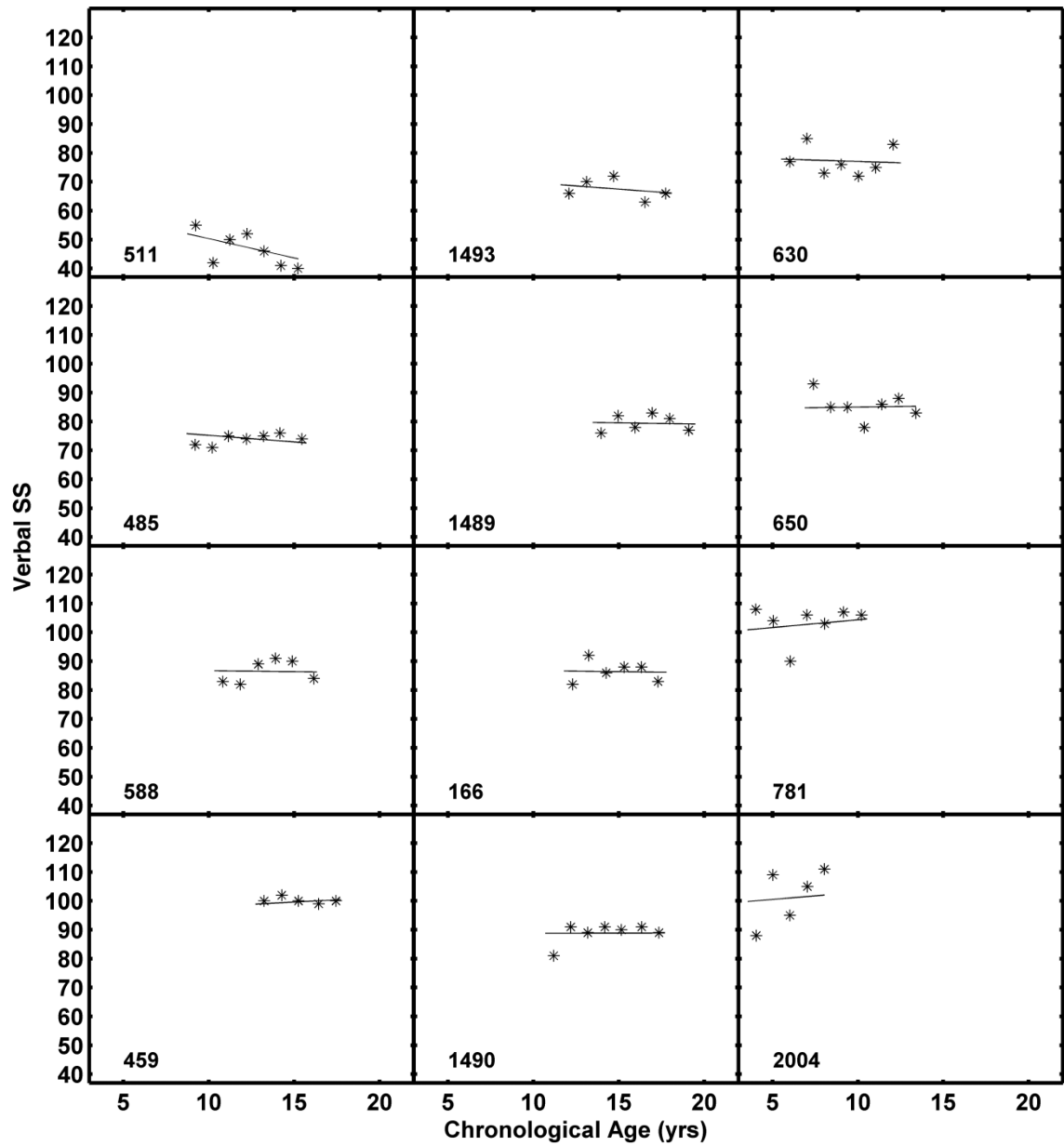


Figure 4. KBIT-2 Verbal standard scores (SS) with fitted functions for the 12 children shown in Figure 2.

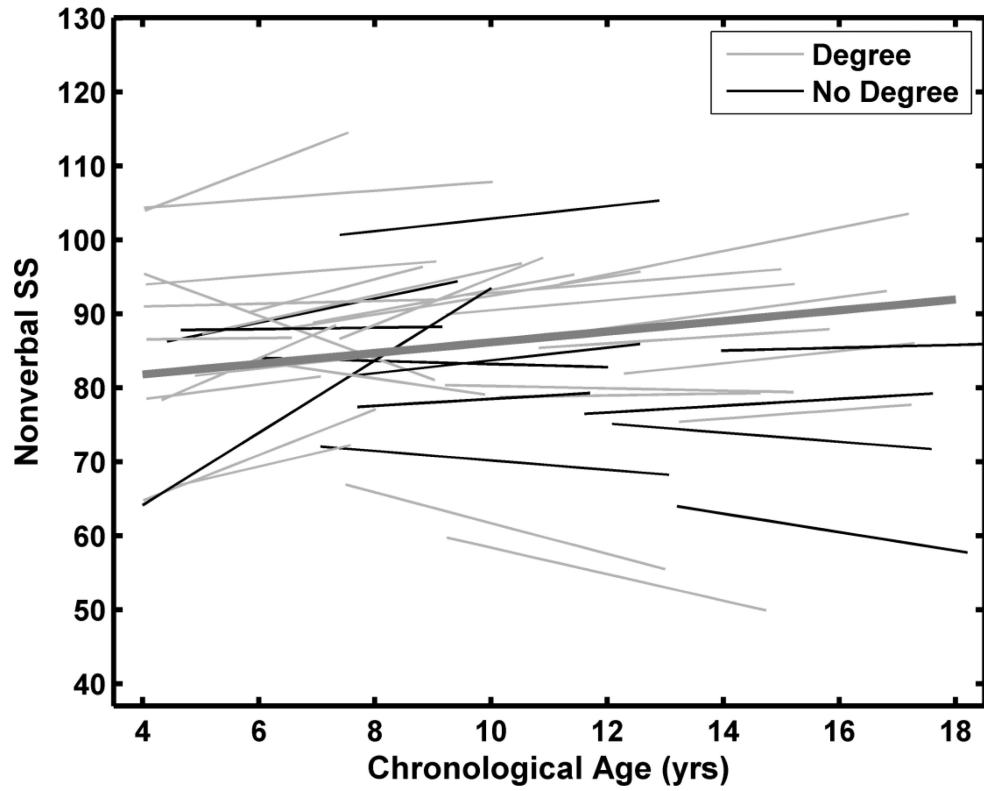


Figure 5. Predicted average developmental trajectories and individual fitted functions for the KBIT-2 Nonverbal standard scores (SS). The function fits were based on the unconditional model. No degree = High school diploma or some college. Degree = Bachelor or more advanced degree.

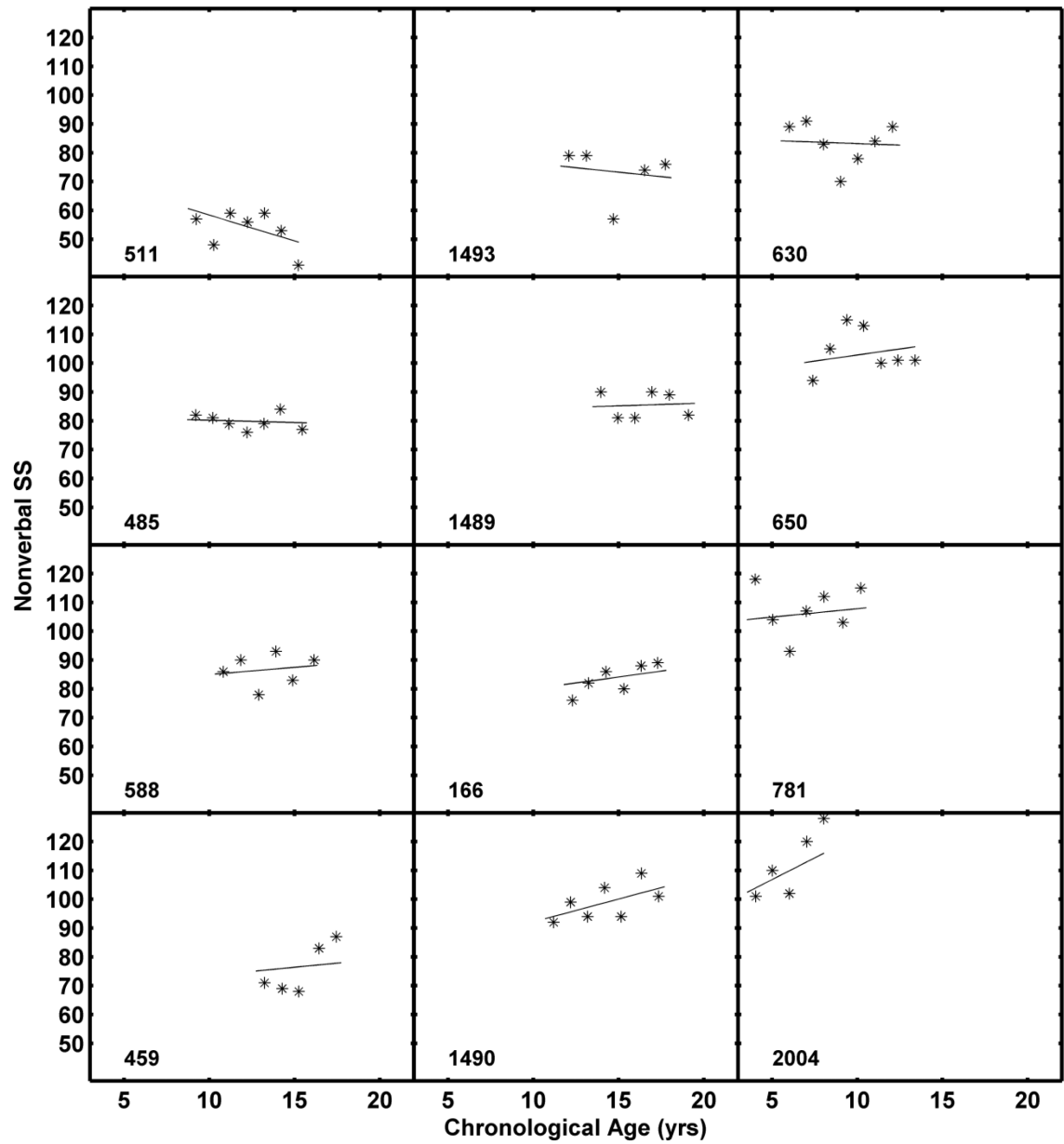


Figure 6. KBIT-2 Nonverbal standard scores (SS) with fitted functions for the 12 children shown in Figure 2.

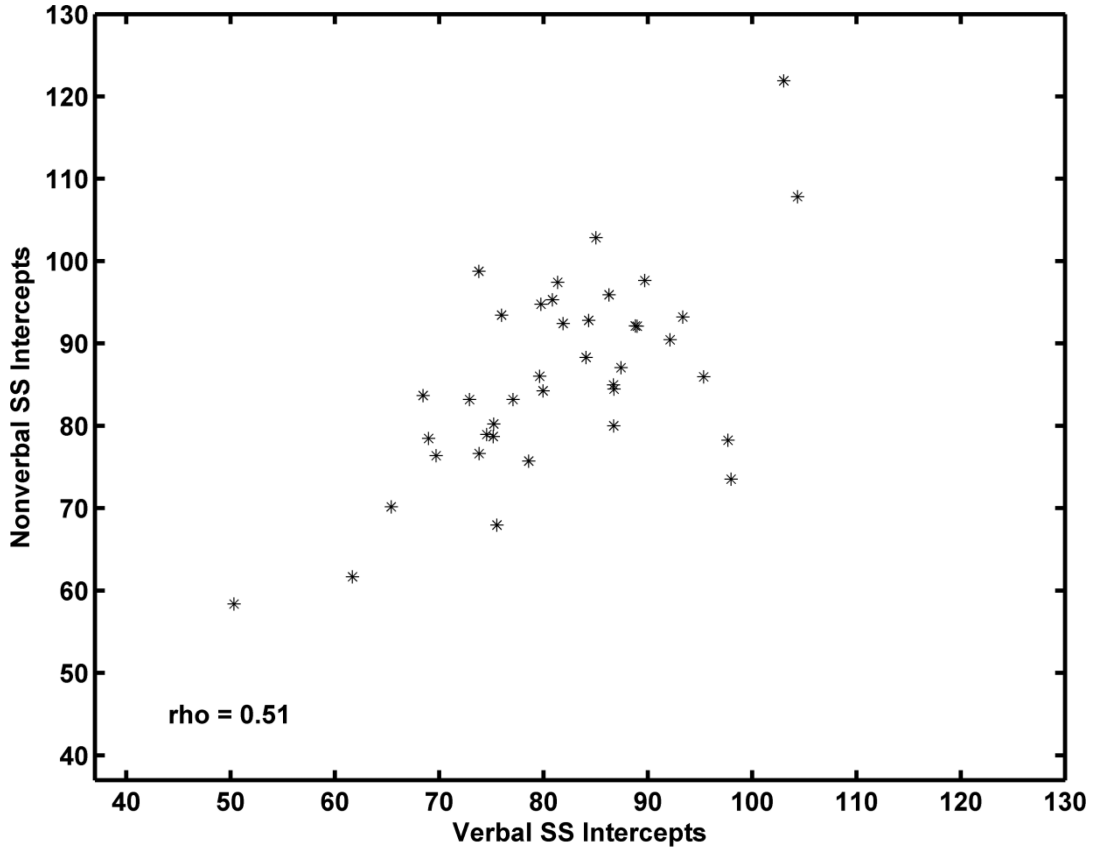


Figure 7. Multilevel model estimates of Nonverbal standard score (SS) intercepts (estimated age 10 scores) plotted as a function of Verbal standard score (SS) intercept estimates. The Spearman correlation is given in the lower left corner.

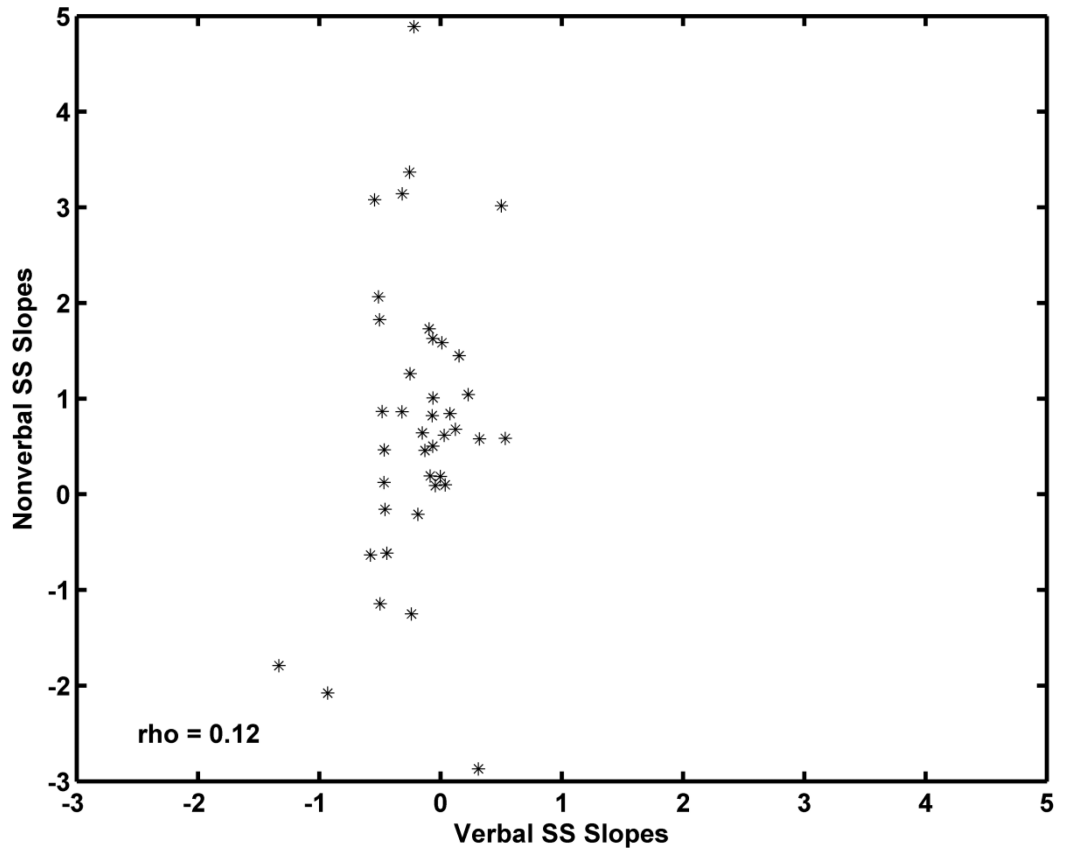


Figure 8. Multilevel model estimates of Nonverbal standard score (SS) slopes plotted as a function of Verbal standard score (SS) slope estimates. The Spearman correlation is given in the lower left corner.

Table 1

Prior longitudinal studies of IQ in individuals with Williams syndrome.

Study (N)	Time 1			Time 2		
	Measure	Mean CA (SD) ^a	Mean SS (SD)	Measure	Mean CA (SD) ^a	Mean SS (SD)
Crisco, 1990 (14)	S-B form L-M IQ	4.2 yrs	67.00 (16.53)	S-B form L-M IQ	9.26 yrs	66.00 (11.88)
Fisch et al., 2010 (18)	S-B-4 IQ	8.85 yrs (3.90)	51.65 (11.55)	S-B-4 IQ	10.85	49.59 (6.50)
Gosch & Pankau, 1996 (18)	CMMS IQ	6.5 yrs	77 (14.7)	CMMS IQ	8.5 yrs	68 (13.1)
	DAP IQ	6.5 yrs	64 (13.2)	DAP IQ	8.5 yrs	65 (10.7)
Udwin et al., 1996 (23)	WISC-R full-scale IQ	12.92 yrs (1.91)	49.5 (11.1)	WAIS-R full scale IQ	21.75 yrs (1.90)	60.8 (5.6)
	WISC-R Verbal IQ		55.7 (10.9)	WAIS-R Verbal IQ		64.2 (5.9)
	WISC-R Perform. IQ		52.5 (9.1)	WAIS-R Perform. IQ		59.7 (4.9)
Howlin et al., 2010 (47)	WAIS-R full-scale IQ	25.75 yrs	62.1 (6.5)	WAIS-III full scale IQ	36.83 yrs	58.2 (7.5)
	WAIS-R Verbal IQ		65.8 (6.9)	WAIS-III Verbal IQ		62.9 (7.8)
	WAIS-R Perform. IQ		61.5 (5.4)	WAIS-III Perform. IQ		60.1 (6.0)

Note. CA = chronological age, SD = standard deviation, SS = standard score, S-B form L-M = Stanford-Binet Intelligence Scale Form L-M (Terman & Merrill, 1972), S-B-4 = Stanford-Binet Intelligence Scale 4th edition (Thorndike, Hagen, & Sattler, 1986), CMMS = Columbia Mental Maturity Scale, German edition (Bondy, Cohen, Eggert, & Luer, 1969), DAP = Draw-A-Person Test (Zaier, 1971), WISC-R = Wechsler Intelligence Scale for Children-Revised (Wechsler, 1976), Wechsler Adult Intelligence Scale-Revised (Wechsler, 1981), WAIS-III = Wechsler Adult Intelligence Scale-III (Wechsler, 1997). For the Stanford Binet, mean IQ = 100 and SD = 16 for the general population; for all other measures included in Table 1, mean IQ = 100 and SD = 15 for the general population.

^aIf available.

Table 2

Prior studies of effect of maternal education level on child IQ.

Study	Design (N)	Group	CA	Measure	Findings for High Versus Low Maternal Education			
					Full scale IQ	Verbal IQ	Perform. IQ	Additional
Breslau et al. (2001)	Long. (717)	TD	6 yrs 11 yrs	WISC-R WISC-R	Hi > Med > Lo Hi > Med > Lo	-- --	-- --	No effect of mat educ. on IQ change from 6 – 11 years
Jedrychowski et al. (in press)	C-S (468)	TD	7 yrs	WISC-R	Hi > Lo	--	--	
Klebanov & Brooks-Gunn (2006)	Long. (228)	TD	5 yrs 8 yrs	WPPSI WISC-R	ns Hi > Lo	-- --	-- --	
Cornelius et al. (2010)	Long. (290)	TD	6 yrs 10 yrs	SB-4 SB-4	Hi > Lo Hi > Lo	-- --	-- --	No effect of mat educ. on IQ change from 6 – 10 years
Bee et al. (1982)	Long. (169)	TD	36 mos 48 mos	SICD SB-LM	-- Hi > Lo	Hi > Lo --	----	
Gale et al. (2004)	C-S (221)	TD	9 yrs	WASI	Hi > Lo	Hi > Lo	Hi > Lo	Effect of mat educ. is larger for Verbal IQ than Performance IQ
Leversen et al. (2011)	C-S (194) C-S (54)	GA 22-27 wks GA 28-32 wks	5 yrs 5 yrs	WPPSI-R WPPSI-R	Hi > Lo ns	-- --	-- --	
Saavalainen et al. (2006)	C-S (40)	preterm	16 yrs	WAI-R	--	Hi > Lo	--	
Ment et al. (2003)	Long. (296)	VLBW	4.5 yrs 6 yrs 8 yrs	WPPSI-R WPPSI-R WISC-III	-- -- --	Hi > Lo Hi > Lo Hi > Lo	-- -- --	
Vohr et al. (2003)	C-S (328)	VLBW	8 yrs	WISC-III	Hi > Lo	Hi > Lo	Hi > Lo	Largest effect is for Verbal IQ
Sommerfelt et al. (1995)	C-S (262)	LBW (116) combined with TD (146)	5 yrs	WPPSI-R	--	Mat educ. contributes significantly to Verbal IQ	Mat educ. not significant contributor to Perform. IQ	

Note: Long. = longitudinal, C-S = cross-sectional, TD = typically-developing, GA = gestational age, VLBW = very low birth weight, LBW = low birth weight, yrs = years, mos = months, WISC-R = Wechsler Intelligence Scale for Children-Revised (Wechsler, 1976), WPPSI = Wechsler Preschool and Primary Scale of Intelligence (Wechsler, 1967), SB-4 = Stanford-Binet Intelligence Scale 4th edition (Thorndike, Hagen, & Sattler, 1986), SICD = Sequenced Inventory of Communication Development (Hedrick, Prather, & Tobin, 1975), SB-LM = Stanford-Binet Scale of Intelligence form L-M (Terman & Merrill, 1972), WASI = Wechsler Abbreviated Scale of Intelligence (Wechsler, 1999), WAI-R = Wechsler Adult Intelligence Scale-Revised (Wechsler, 1981), WISC-III = Wechsler Intelligence Scale for Children-III (Wechsler, 1991), hi = high maternal education group, med = medium maternal education group, lo = low maternal education group, ns = not significant, mat educ. = maternal education level, Perform. IQ = Performance IQ.

Table 3

Multilevel model statistics for KBIT-2 Composite IQ unconditional model.

Fixed Effects	Coefficient	Standard Error	t Ratio (df)	<i>p</i>
Intercept (SS at age 10) π_0				
Mean Intercept β_{00}	-18.95	2.10	-9.02 (39)	<.0001
Slope (SS annual change) π_1				
Mean Slope β_{10}	0.37	0.26	1.44 (39)	0.16
Random Effects				
	Standard Deviation	Variance	X^2 (df)	<i>p</i>
Intercept r_0	12.83	164.75	624.55 (39)	<.0001
Slope r_1	1.04	1.08	68.14 (39)	0.003
Level 1 (Within-person) e	5.29	27.96		

Note. Intercept estimates are deviations from the KBIT-2 normed mean of 100.

Table 4

Multilevel model statistics for KBIT-2 Composite IQ maternal education model.

Fixed Effects	Coefficient	Standard Error	<i>t</i> Ratio (<i>df</i>)	<i>p</i>
Intercept (SS at age 10) π_0				
No degree mean β_{00}	-24.44	3.74	-6.53 (38)	<.0001
College deg. mean difference β_{01}	7.85	4.48	1.75 (38)	0.09
Slope (annual change) π_1				
No degree mean β_{10}	0.31	0.44	0.69 (38)	0.50
College deg. mean difference β_{11}	0.11	0.55	0.20 (38)	0.84
Random Effects				
	Standard Deviation	Variance	χ^2 (<i>df</i>)	<i>p</i>
Intercept r_0	12.54	157.26	595.33 (38)	<.0001
Slope r_1	1.06	1.13	68.24 (38)	0.002
Level 1 (Within-person) e	5.29	27.96		

Table 5

Multilevel model statistics for KBIT-2 Verbal standard score unconditional model.

Fixed Effects	Coefficient	Standard Error	<i>t</i> Ratio (<i>df</i>)	<i>p</i>
Intercept (SS at age 10) π_0				
Mean Intercept β_{00}	-18.46	1.82	-10.16 (39)	<.0001
Slope (SS annual change) π_1				
Mean Slope β_{10}	-0.19	0.18	-1.08 (39)	0.29
Random Effects				
	Standard Deviation	Variance	X^2 (<i>df</i>)	<i>p</i>
Intercept r_0	11.30	127.71	628.50 (39)	<.0001
Slope r_1	0.37	0.14	44.18 (39)	0.26
Level 1 (Within-person) e	4.73	22.36		

Table 6

Multilevel model statistics for KBIT-2 Verbal standard score maternal education model.

Fixed Effects	Coefficient	Standard Error	<i>t</i> Ratio (<i>df</i>)	<i>p</i>
Intercept (SS at age 10) π_0				
No degree mean β_{00}	-24.19	3.18	-7.60 (38)	<.0001
College deg. mean difference β_{01}	8.12	3.81	2.13 (38)	0.04
Slope (annual change) π_1				
No degree mean β_{10}	-0.23	0.30	-0.76 (38)	0.45
College deg. mean difference β_{11}	0.06	0.38	0.17 (38)	0.87
Random Effects				
	Standard Deviation	Variance	χ^2 (<i>df</i>)	<i>p</i>
Intercept r_0	10.82	117.13	547.68 (38)	<.0001
Slope r_1	0.38	0.15	44.01 (38)	0.23
Level 1 (Within-person) e	4.74	22.43		

Table 7

Multilevel model statistics for KBIT-2 Nonverbal standard score unconditional model.

Fixed Effects	Coefficient	Standard Error	<i>t</i> Ratio (<i>df</i>)	<i>p</i>
Intercept (SS at age 10) π_0				
Mean Intercept β_{00}	-13.87	2.21	-6.23 (39)	<.0001
Slope (SS annual change) π_1				
Mean Slope β_{10}	0.72	0.40	1.82 (39)	0.08
Random Effects				
	Standard Deviation	Variance	X^2 (<i>df</i>)	<i>p</i>
Intercept r_0	13.00	169.06	359.51 (39)	<.0001
Slope r_1	1.93	3.71	92.34 (39)	<.0001
Level 1 (Within-person) e	7.02	49.34		

Table 8

Multilevel model statistics for KBIT-2 Nonverbal standard score maternal education model.

Fixed Effects	Coefficient	Standard Error	<i>t</i> Ratio (<i>df</i>)	<i>p</i>
Intercept (SS at age 10) π_0				
No degree mean β_{00}	-17.06	4.00	-4.26 (38)	<.0001
College deg. mean difference β_{01}	4.61	4.81	0.96 (38)	0.34
Slope (annual change) π_1				
No degree mean β_{10}	0.63	0.71	0.89 (38)	0.37
College deg. mean difference β_{11}	0.17	0.86	0.20 (38)	0.84
Random Effects				
	Standard Deviation	Variance	χ^2 (<i>df</i>)	<i>p</i>
Intercept r_0	13.10	171.60	350.21 (38)	<.0001
Slope r_1	1.97	3.87	92.55 (38)	<.0001
Level 1 (Within-person) e	7.02	49.31		