Number line estimation from kindergarten to grade 2: a longitudinal study

Praet, M., & Desoete, A. (Learning and Instruction, in press)

Number line estimation from kindergarten to grade 2: a longitudinal study

Abstract

A bulk of evidence supports the association of number line estimations using Arabic digits and dots with math learning. Surprisingly few studies have been conducted to explore the relationship between estimations using number words and mathematics. The present study expands previous findings by investigating estimations in three formats (Arabic digits, dots and number), adding language as predictor and by focusing at timed and untimed math learning. A sample of 132 children was followed from kindergarten till grade 2. Results reveal variability in estimation accuracy and errors declining with age and instruction in all children. In addition, our findings suggest that Arabic numerals have a more linear distribution than number words. Moreover, our findings suggest that language explains variation in kindergarten but not in evolution and, more in particular, untimed math achievement can be predicted by number line estimation. Implications for assessment, prediction of math learning and instruction are discussed.

Keywords: Number line, estimation accuracy, representation, distribution, growth curve, language, number words, inversion, untimed mathematic achievement, timed arithmetic fluency.

Highlights

° Variability in estimation accuracy but not in evolution.

° Arabic numerals have a more linear distribution than number words.

° Number words dominant as format in grade 2.

° Language explains variation in kindergarten but not evolution.

° Untimed math learning can be predicted by estimation skills

1. Introduction

There has been extensive research on number line estimation (Berteletti, Lucangeli, Piazza, Dehaene, & Zorzi, 2010; Schneider et al., 2008; Slusser, Santiago, & Barth, 2013) and the relationship with mathematics. However, most of these studies have a cross-sectional design, using dots or Arabic numbers (e.g., Ashcraft & Moore, 2012; Geary, Hoard, Nugent, & Byrd-Craven, 2008; Moeller, Pixner, Kaufmann, & Nuerk, 2009; Muldoon, Towse, Simms, Perra, & Menzies Towse, 2013; Sasanguie, Göbel, Moll, Smets, & Reynvoet, 2013) as stimuli for the estimations. In addition to estimation, the value of including language as predictor for mathematics has been stressed (Praet, Titeca, Ceulemans, & Desoete, 2013; Sarnecka et al., 2007; Wiese, 2003). However surprisingly few studies have been conducted to explore the combined effect of these predictors on math learning. This study addresses this gap and has the unique scientific merit of focusing on age-related changes in children's numerical estimate accuracy and distribution using three different format types (stimuli as Arabic numerals, spoken number words, and dot patterns) at five measuring points (from kindergarten to Grade 2), with children becoming more familiar with numbers up to 100 (learning to count in kindergarten, deal with numbers from 0 to 20 in grade 1 and up to 100 in grade 2). In addition this study expands previous findings, by investigating the prediction for timed and untimed math learning. Insight about the detailed nature of underlying number representations can inform targeted assessment and might have educational implications for learning and instruction researchers and professional addressing kindergarteners at risk for mathematical learning difficulties.

1.1. Numerical estimation and development

It is widely accepted that there is a gain in accuracy of number line judgments on a 0-100 interval with increasing formal schooling. In addition research indicates a developmental transition from a logarithmic distribution of the representation of numbers (with children experiencing a larger distance between 2 and 3 than between 18 and 19) to a more linear function as the result of a better one-to-one correspondence between the value being judged and its estimate, from preschool to primary school (Siegler & Booth, 2004; Siegler & Opfer, 2003). The linearity of judgments is often positively correlated with math learning (Ashcraft & Moore, 2012; Siegler & Booth, 2004).

Dehaene stated in his triple code model that number representation takes place in three different ways, with three different formats, located at three different brain regions (Dehaene, 1992; 1997, Dehaene & Cohen, 1995). Firstly, there is a (symbolic) visual system where numbers are encoded as strings of Arabic digits (e.g., '14') needed for multidigit calculation and parity judgments. Secondly, there is a (symbolic) verbal system where numbers are represented as sequences of number words (e.g., 'fourteen') lexically, phonologically and syntactically. The third system uses (asymbolic) analogue magnitude codes as non-verbal semantic size and distance relations between (e.g., a collection of 14 dots).

Although evidence was found for a general, modality-independent representation across different kind of magnitudes, such as numbers, quantities of objects, lengths and durations (Barth, Kanwisher, & Spelke, 2003; Huntley-Fenner, & Cannon, 2000), some studies have found a relationship between symbolic tasks but not between non-symbolic number comparison skills and math learning (e.g., De Smedt, Noel, Gilmore, & Ansari, 2013; Holloway & Ansari, 2009; Mundy & Gilmore, 2009). In addition, up to now, most studies focussed on non-symbolic magnitude representation, sometimes in combination with the symbolic representation with Arabic numbers (e.g., Ashcraft & Moore, 2012; Geary et al., 2008; Moeller et al., 2009; Muldoon et al., 2013; Sasanguie et al., 2013). On basis of such data, it is often unclear whether it is the Arabic number or number words processing that is important for math learning. Finally, Sasanguie and colleagues (2013) suggested an association between estimation and a general curriculum-based math test but not with a timed math fluency test. Therefore we might question ourselves if the used format to test number estimation or math learning does not affect an influence on the observed relationships.

The last decades, several researchers have studied the relationship between estimation and mathematics achievement . Recently, Muldoon and colleagues (2013) revealed in a longitudinal study that 5-year olds with less-accurate internal representations of numbers tested on 4 occasions at 3 months intervals en were disadvantaged on some early math tasks, such as recognizing number names and numerals, identifying quantitative relationships, matching magnitudes and quantities or solving easy word problems, compared to peers with better quality representations. However, the question of whether it is the Arabic number or number words processing that is important for math learning and the relationship with math fluency and untimed math learning remains unresolved. Since the other existing research are cross-sectional studies (Ashcraft & Moore, 2013; Berteletti, Lucangeli, & Zorzi, 2012; Booth & Siegler, 2006; Ebersbach, Luwel, Frick, Onghena, & Verschaffel, 2008; Holloway & Ansari, 2009; Sasanguie et al., 2013; Schneider et al., 2008; Siegler & Booth, 2004; Slusser et al., 2013), predictions on estimation accuracy and distribution growth are difficult to make.

1.2. Math achievement and language

Although children process numbers long before the acquisition of language (Dehaene, 2001), the value of including language has recently been stressed in the prediction of numeracy development (Praet et al., 2013; Purpura, Hume, Sims, & Lonigan, 2011; Romano, Babchishin, Pagani, & Kohen, 2010; Sarnecka et al., 2007; Wiese, 2003).

Having a larger nominal vocabulary was found to be helpful in the acquisition of number words (Negen & Sarnecka, 2012). In addition some studies (Barner, Chow, & Yang, 2009; Negen & Sarnecka, 2012) revealed that general measures of language development also predicted number-word knowledge, although other studies (e.g., Ansari, Donlan, Thomas, Ewing, Peen, & Karmiloff-Smith, 2003) did not find such a link. Whether or not language helps children in kindergarten to solve mathematical problems, remains a point of discussion.

1.3. The current study

To summarize, empirical evidence for age-related changes in estimations in three formats, adding language as predictor and focusing at timed and untimed mathematic achievement is lacking. Moreover, very few studies examined these skills in a longitudinal design from kindergarten till grade 2.

This study addresses the following two major research questions: (a) is the accuracy and distribution of the estimation of the position of numbers using different formats (stimuli as Arabic numerals, spoken number words, and dot patterns) mirroring the familiarity with numbers and predicting untimed and timed math learning? And (b) Does language explains variation in the growth curves?

For the first research question four additional questions or hypotheses were formulated. We expected a better accuracy in the estimation of the position of numbers in older children mirroring their familiarity with numbers (Hypothesis 1). Considering the format-independency hypothesis, similar results on the estimation with Arabic numerals, spoken number words, and dot patterns were expected (Hypothesis 2). In line with the developmental shift, we expected a kindergarteners estimating in a logarithmic manner, and children in grade 1 and grade 2 following a more linear curve (Hypothesis 3). Finally different predictions for the processing of untimed calculation and timed fact retrieval tasks are expected (Hypothesis 4).

For the second research question changes over time were expected with language explaining some of the variation in the growth curves (Hypothesis 5).

2. Methods

2.1. Participants and procedure

The children in this study (*N*=132, 53% girls) were Dutch-speaking children from five kindergartens serving children from families with working and middle-class-socio-economic backgrounds. Written parental consent to participate in the study was obtained for all children.

All children were individually tested in kindergarten at measurement time 1 (March kindergarten = time period T1) in a quiet room of the school to obtain measures of intelligence, number estimation and early calculation skills.

Measurement 2 and 3 took place in grade1 (November grade 1 = time period T2 getting instruction on numbers 0-10, June grade 1 = time period T 3 getting instruction on numbers 0-20). All children were individually tested on their number estimation and ability to solve simple calculations (T2 and T3) as well as on their ability to retrieve number facts (T3).

Due to constraints in access to schools and the children attending them, it was not feasible to collect on all children on all five time points. On half of the children data were collected on all five time points. They were tested in grade 2 (October grade 2 = time period

T4 rehearsal of instruction on numbers 0-20, January grade 2 = time period T5 getting instruction on numbers 0-100) on number estimation (T4 and T5) and on their ability to calculate (T5).

2.2. Measures

2.2.1. Intelligence

Intelligence was assessed in kindergarten (at T1) with the Wechsler Preschool and Primary Scale of Intelligence or the WPPSI-III-NL (Wechsler et al., 2002; Hendriksen & Hurks, 2009). Children completed the three core verbal tests (information, vocabulary, and word reasoning) and the three performal tests (block patterns, Matrix reasoning, and concepts drawing).

2.2.2. Language skills

To get a picture of the oral language skills in kindergarten (at T1) all the children were tested with the Clinical Evaluation of Language Fundamentals or the CELF-$_4$Nl (Semel, Wiig, & Secord 2008; Kort, Schittekatte, & Compaan 2008). This resulted in a core language score. This test is validated on 1280 children. The internal consistency was good, with Cronbach's alpha between .87 and .95.

2.2.3. Number estimation skills

All children were tested with a forced choice number line estimation task in kindergarten (T1 at the age of *M*=68 months, *SD*=4 months), November grade 1 (T2) and June grade 1 (T3). Half of them were followed up in grade 2 (October grade 2 T4, January grade 2 T5).

The ***Number Line Estimation*** (NLE) task used, in line with Berteletti and colleagues (2010) and Booth and Siegler (2006) a 0-100 scale. The task included three exercise trials and 30 test trials. Stimuli (2, 3, 4, 6, 18, 25, 42, 67, 71 and 86) were presented in three different formats, as Arabic numerals not read aloud to the children (e.g. anchors 0 and 100, target number 25), spoken number words written on the scale (e.g. anchors zero and hundred, target number twenty-five), and dot patterns (e.g. anchors of zero dots and hundred dots, target number twenty five dots). A higher proportion of smaller numbers (2, 3, 4, 6, 18) compared to larger numbers (25, 42, 67, 71 and 86) was used to obtain a fine-grained data pattern for the lower number range with children in grade 1 only knowing numbers up to 20. Magnitude estimations were compared in three formats. The Arabic numeral and number word estimation tasks were symbolic estimation tasks, with subjects having to make a numerical translation from the assignment of Arabic numerals or number words to its position on a line. The dot estimation task was a magnitude estimation task, with subjects having to estimate a quantity by indication its position on a line with dots as anchors. The dot patterns were controlled for perceptual variables using the procedure of Dehaene, Izard and Piazza (2005), meaning that on half of the trials dot size was held constant, and on the other half, the size of the total occupied area of the dots was held constant. Children were asked to put a single mark on the line to indicate the location of the number (Berteletti et al., 2010): “We will now play a game with numbers. Look at this page, you can see a long line, ranging from zero to ten. Above the line, you can see a number/the number x/ dots. I want you to show me where this number/the number x/the dots should be on the line. If here is zero, and here is ten, were should this number/the number x/these dots be located on the line? If you know where this number/ the number x/ these dots belong, you can make a single mark with your pencil on the line.” No feedback was given to participants regarding the accuracy of their marks. The instructions could be rephrased if needed, but no suggestions were given on the correct place of the mark. The percentage absolute error (PAE) – the amount by which their estimated deviated from the correct values-was calculated per child as a measure of children’s estimation accuracy following formula by Siegler and Booth (2004).

2.2.4. Mathematics achievement

To assess math learning, outcome measures were used focusing on what the children are supposed to have learned during formal math education according to their grade curriculum.

To assess kindergarteners’ skills (at T1) subtest five of the **Tedi-Math** (Grégoire, Noël, & Van Nieuwenhoven, 2004) was used. This untimed subtest consists of simple arithmetic operations to measure early numeracy in kindergarten. The child was presented simple arithmetic operations on pictures (e.g. ‘Here you see two red balloons and three blue balloons. How many balloons are there together?’). Cronbach’s alpha was .84. The Tedi-Math was used and tested for conceptual accuracy and clinical relevance in previous studies (e.g., Stock, Desoete, & Roeyers, 2010).

At T2 and T5 children completed the **Kortrijk Arithmetic Test Revision** (KRT-R; Baudonck et al., 2006) as untimed general curriculum-based mathematics achievement test. The KRT-R is a standardized test which requires that children solve 30 simple calculations in a number-problem format (e.g., 16 - 12 = …), and 30 more complex calculations often in a word-problem format (e.g., 1 less than 8 is …). The test focuses on what children are supposed to have learned on number knowledge, mental arithmetic and procedural calculation according to their grade curriculum. The test is different for grade 1 (T2) and grade 2 (T5). Thus, the same constructs are included in each grade but at a different difficulty level. The psychometric value of the test has been demonstrated on a sample of 3,246 children. The validity coefficient (correlation with school results) varies between .64 and .66 and reliability coefficient (Cronbach's alpha) between .83 and .94 indicating good psychometric values.

At T3 children completed the ***CDR Test*** (Desoete & Roeyers, 2006), an untimed curriculum-based standardized test on simple calculations in a number-problem format (e.g., 16 - 12 = …), or in a word-problem format (e.g., 1 less than 8 is …). The psychometric value of the test has been demonstrated on a sample of 1,792 children. The reliability coefficient (Cronbach's alpha) was .93.

At T3 and T5 children also completed the **Arithmetic Number Fact Test** (Tempo Test Rekenen (TTR; De Vos, 1992), a timed arithmetic test, to assess their fact retrieval skills. The TTR (De Vos, 1992) is a test consisting of 80 (first grade) or 200 (second grade) arithmetic number fact problems. In first grade children have to solve as many additions and subtractions in two minutes, children in the second halve of second grade are presented additions, subtractions, divisions and multiplications and have five minutes to solve as many as possible items. The TTR is a standardized test that is frequently used in Flemish education as measure of early arithmetic acquisition. The total number of correct items was used as score for the analyzes. The psychometric value of the TTR has been demonstrated on a sample of 10,059 children in total. The test is identical for period T3 and T5. For the analyses the total number of correct items was used.

2.3. Analysis procedure

Before testing the hypotheses, the PAEs were log transformed for distributional reasons.

The mean PAE was be analysed in kindergarten (T1), grade 1 (T2 and T3) and grade 2 (T4 and T5). In addition, to study the shift from logarithmic to linear representation and the relationship with arithmetic, a number of regression analyses were conducted (in kindergarten, grade 1 and 2). Moreover the prediction for dots was compared with the predictive value of symbolic stimuli (Arabic digits and number words) simultaneously entered as predictors in the five time periods. In addition, regression analyses were conducted to study cross-sectional relationships between PAE and arithmetic measures in kindergarten (T1), grade 1 (T2 and T3) and grade 2 (T5). Finally, it was explored whether the R^2_{lin} values of linear fits could predict the arithmetic achievement at T1,T2, T3 and T5.

A latent growth curve model was fitted with the intercept as logPAE (accuracy level) and the slope as linear growth rate. The growth model was used to study the changes in relationships between the variables over time. Unnested models were compared based on Akaike information criterion (AIC). This AIC is a measure of the relative quality of a statistical model, for a given set of data. AIC deals with the trade-off between the complexity of the model and the goodness of fit of the model. The lower the AIC, the better the relative quality of the statistical model. In addition the following 'goodness of fit' indices were reported: relative chi-square (χ^2/df) attempting to make the index less dependent on the model complexity, the Comparative Fit Index (CFI) and the Root Mean Square Error or Approximation (RMSEA) not requiring comparison with a null model.

3. Results

3.1. Descriptive statistics

IQ and language were assessed in kindergarten. Children had an average Total Intelligence (TIQ; *M* = 101.39, *SD* =12.73).; Verbal Intelligence (VIQ; *M* = 102.74; *SD* =11.97) and Performance Intelligence (PIQ; *M* = 99.29, *SD* =11.68) were average assessed with the WPPSI. The language core index on the CELF-IV was 98.17 (*SD* = 11.40).

3.2.Numerical estimation and development

3.2.1. Estimation accuracy : development from kindergarten till the middle of grade 2

The magnitude representation inaccuracy on the total test or the Percentage of Absolute Error (PAE) in the estimation on all 30 trials (format independent) on the number line task and the results on the different formats from kindergarten to Grade 2 is described in Table 1.

< Insert Table 1 about here >

Table 1 reveals that the overall estimations become more accurate when children get older and more familiar with numbers. The PAE decreased 6.51% from kindergarten to the beginning of grade 1. The PAE only decreased 2.26% from the end of grade 1 to October of grade 2, making the decrease slow down . However, children made less errors (PAE decreased) from kindergarten to grade 2.

3.2.2. Estimation accuracy: format-indepency

Pairwise comparisons between of the PAE on **Arabic numerals** revealed significant differences between time period 1 and 2 ($p < .001$) , time 2 and 3 ($p < .001$), time 3 and 4 ($p = .002$) but not between time 4 and 5 ($p = .239$).

The pairwise comparisons of the PAE on **number words** revealed no significant differences between time 1 and 2 ($p < .603$) , nor between time 2 and 3 ($p = .215$), a significant difference between time 3 and 4 ($p = .012$), and time 4 and 5 ($p = .024$).

Pairwise comparisons of the PAE on **dots** revealed no significant differences between time 1 and 2 ($p = .548$) , time 2 and 3 ($p = .272$), a significant difference between time 3 and 4 ($p = .008$), but not between time 4 and 5 ($p = .682$).

Thus, format-independent, the estimations become more accurate from kindergarten (period 1) to the end of grade 1 (period 3) on all estimation tasks (using Arabic numerals, number words or dots as formats). In addition, format-independent there was no significant difference between the estimation accuracy at the beginning (period 4) and middle (period 5) of grade 2.

3.2.3. **Estimation : developmental shift in the distribution**

The first regression analyses were conducted on **group level** on all 30 trials (in kindergarten, grade 1 and 2) revealed a significant **logarithmic representation** (t (129) = 12.712, $p < .001$, $R^2_{log} = .959$, $R^2_{lin} = .729$) **in kindergarten** for the total number line test (see Figure 1). The representation with Arabic numbers (t (129) = 8.561, $p < .001$, $R^2_{log} = .946$, $R^2_{lin} = .751$), number words (t (129) = 5.935, $p < .001$, $R^2_{log} = .968$, $R^2_{lin} = .766$) and dots (t (129) = 8.304, $p < .001$, $R^2_{log} = .927$, $R^2_{lin} = .740$) also had a logarithmic distribution.

< Insert Figure 1 about here >

In November of **grade 1** children had a significant **logarithmic** representation of numbers on the total test (*t* (129) = 8.096, p < .001, R^2_{log} = .970, R^2_{lin} = .843), see Figure 2. The representation with Arabic numbers (*t* (129) = 8.777, p < .001, R^2_{log} = .976, R^2_{lin} = .848) , number words (t (129) = 6.497, p < .001, R^2_{log} = .984, R^2_{lin} = .894) and dots (*t* (129) = 4.112, p < .001, R^2_{log} = .751, R^2_{lin} = .737) also had a logarithmic distribution.

< Insert Figure 2 about here >

At the end of grade 1 children had a significant **logarithmic** representation on the total test (*t* (126) =5.962, p< .001, R^2_{log} = .961, R^2_{lin} =.899), see Figure 3. The representation with Arabic numbers (*t* (126) = 5.456, p < .001,R^2_{log} = 969., R^2_{lin} =.911), number words (*t* (126) = 5.595, p < .001,R^2_{log} = .976, R^2_{lin} = .917) and dots (t (126) = 3.070, p = .003,R^2_{log} = .854, R^2_{lin} =.833) also had a logarithmic distribution

< Insert Figure 3 about here >

At the **beginning** of **grade 2** children still had a significant logarithmic representation on the total test (*t* (58) =2.102, p = .040, R^2_{log} = .956, R^2_{lin} =.938), see Figure 4. The representation with number words (*t* (58) = 2.602, p = .012,R^2_{log} = .955, R^2_{lin} =.952) and dots (t (58) = 2.297, p = .025,R^2_{log} = .894, R^2_{lin} =.914) also had a logarithmic distribution. This was not the case for Arabic numbers (*t* (58) = 0.034, p = .973), R^2_{log} = .933, R^2_{lin} =.965

< Insert Figure 4 about here >

In the **middle** of grade 2 there was no longer a significant logarithmic representations on the total test (*t* (61) =0.997, p = .323, R^2_{log} = .892, R^2_{lin} =.977 - see Figure 5). In addition, there was a significant **linear distribution** for Arabic numbers (*t* (61) = 2.701, p = .009,R^2_{log} = .888, R^2_{lin} =.992), number words (*t* (61) = 2.953, p = .004,R^2_{log} = .890, R^2_{lin} =.992) and dots (*t* (61) = 2.029, p = .047,R^2_{log} = .863, R^2_{lin} =.932).

< Insert Figure 5 and Table 2 about here >

Because of the longitudinal design it was also possible to look at a more **individual level** to the evolution of the number of children having a linear representation over time (see Table 3).

< Insert Table 3 about here >

Analyses revealed that in kindergarten most children had a logarithmic or no valid representation on the number-to-position task of the 0-100 number line using symbolic stimuli (Arabic numbers, number words) or non-symbolic (dots) stimuli. At the start of grade 1, when children got instruction on numbers 0-10 almost all children had a valid but logarithmic representation for Arabic numbers and number words. At the end of grade 1 with children becoming familiar with numbers up to 20, about one third of them had a linear representation, whereas two third still had a more logarithmic representation of numbers 0-100 on all formats. At the start of grade 2 more than half of the children had a linear representation of Arabic numbers whereas this was only the case for 37.3% and 39% of the representation with number words or dots as stimuli. In the middle of grade 2 the representation became linear for nearly 60% of the children on all formats.

3.2.4. **Estimation : relationship with math learning**

First, the relationship between estimation **accuracy** and math learning was analysed (see Table 4). The cross-sectional relationship between **early mathematics** in **kindergarten** was significant (F (3, 128) = 11.966, p < .001, R^2 = .223) for PAE Arabic digits (p = .013), but not for PAE number words (p = .384) nor for PAE dots (p = .900).

< Insert Table 4 about here >

The regression analysis at the start of **grade 1** on **untimed calculation** skills (F (3, 121) =

4.676, p = .004, R^2 = .326) revealed a trend for PAE with dots (p =.102) but not for Arabic numerals (p = .315) or number words (p =.783; see Table 3). The regression analysis was not significant for **timed fact retrieval** (F (3, 123) = 2.473, p = .065, R^2 = .058).

The regression analysis (on half of the children) in the **middle of grade 2** was significant for **untimed calculation** (F (3, 62) = 7.560, p = .000, R^2 = .278), especially for PAE of number words (p =. 048). The prediction for **timed fact retrieval** was not significant (F (3, 54) = 2.586, p < .063, R^2 = .132).

In addition, we explored whether the linearity of the distribution (R^2_{lin} values of linear fits) could predict math learning. This was marginally the case in kindergarten (F (1, 129) = 3.766, p = .055, R^2 = .029) but not at the beginning of grade 1 (F (1, 125) = 2.347, p = .128, R^2 = .019). At the end of grade 1 the regression analysis was significant for untimed math learning (F (1, 124)=20.758, p < .001, R^2 = .144, but not for timed fact retrieval (F (1, 124) = 3.551, p = .062, R^2 = .028) . Moreover untimed math learning could be predicted by the number line linearity in the middle of grade 2 (F (1, 59) = 18.832, p < .001, R^2 = .245), but this was not the case for timed fact retrieval (F (1, 54) = 1.821, p = .183, R^2 = .033).

3.3. Language and math learning: Growth model

To investigate hypothesis 5, first a latent growth curve model was fitted with logPAE (accuracy level) as outcome variable and random intercept and random slope as linear growth rate. The fit of this model was acceptable based on TLI and CFI (NNFI (TLI) =.878) and CFI = .943 but not acceptable based on $\chi 2$/df= 2.517 and RMSEA = .108 . The estimated intercept for logPAE (mean logPAE at T1) was 3.175 (95% CI: 3.11 to 3.24). The estimated overall change in logPAE (change in logPAE between T 1 and 5) was -0.8 (95%CI: -0.88 to -

0.72). In terms of PAE this meant that the estimated geometric mean for PAE in kindergarten was 23.9 (95% CI: 22.4 to 25.6). The PAE decreased with 54.9% (95% CI: 51.3% to 58.3%) between the end of kindergarten and the middle of grade 2. There was significant interindividual variation for the intercept (estimated *SD*=0.308; $p < .001$) but not for the slope (estimated *SD* = 0.18; Walt test $p = .266$; generalized variance Likelihood ratio test $p = .42$), meaning that there was significant variability between the children on estimation accuracy (intercepts) but not in evolution of their growth curves.

When the different modalities were analyzed, the Wald test ($p =.027$) and generalized variance test (Likelihood ratiotest; $p = .033$) revealed significant variability between the children for the intercepts and slopes using dots as stimuli. However, there was only significant variability for the intercepts but not for the slopes using Arabic numbers (Wald test $p = .497$; Likelihood ratio test $p = .25$) and number words (Wald test $p = .26$; Likelihood ratio test $p = .36$).

In a next step Chi squared was used to compare the model with and without the interaction between slope and IQ (24.217 vs 20.360, df=1; $p = .049$), leading to the choice of a model with IQ x slope interaction and IQ x intercept interaction. Intelligence had a significant effect ($p = .05$) on the inter individual differences on slope, so the interaction was included in the model. The fit of this model was acceptable (NNFI (TLI) =.904) with $\chi 2/df= 2.036$, RMSEA = .089 and CFI= .954.

< Insert Figure 6 about here >

Intelligence explained a significant part ($p <.001$) of the inter individual variability in intercepts of the logPAE. There was also a trend for IQ explaining ($p = .05$) the inter individual variability in the slopes of the logPAE. Standard deviation of the intercepts was reduced from 0.31 to 0.245 by adding IQ as a predictor. Standard deviation of the slopes was only slightly reduced from 0.179 to 0.182 by adding IQ as a predictor. With 1 point increase of IQ, the PAE in kindergarten decreased with 1.5% (95% CI: -1.1% to -1.9%). For an increase of 10 points of IQ, the PAE for T1 is expected to decrease with 13.9% (95% CI: -17.3% to -10.4%). In addition, for one-point increase of IQ, the slope of PAE (PAE at T5-PAE at T1) is expected to increase with 0.5% (95% CI: -1.1% to -1.9%). For an increase of 10 IQ points, the slope of PAE (PAE at T5 - PAE at T1) is expected to decrease with 5.5% (95% CI: 0% to 1.11%).

To investigate whether language could explain some of the inter individual variation in the "growth curves" (hypothesis 5), chi squared was used to compare the model with and without the interaction between language and slope (20.146 vs 19.303 df=1; *p*=.36 leading to the choice of a model without language*slope interaction. The fit of this model (see Figure 10) was good (NNFI (TLI) =.915) with $\chi 2$/df= 1.831, RMSEA = .08, CFI= .956 and AIC = 52.146.

< Insert Figure 7 about here >

Language had a significant effect on the intercept (estimate -.011, S.E. .002, C.R. -4.977, *p* = .002) but not on growth, meaning that for a one-unit increase of the language core index the PAE in kindergarten is expected to decrease with 1.1% (95% CI: 0.7% to 1.5%).

4. Discussion

The importance of predictors for the development of mathematics has been demonstrated (e.g., Kolkman, Kroesbergen, & Leseman, 2013). The current study is the first to simultaneously tap the contribution of number words in addition to estimation using Arabic numbers and dots in the acquisition of timed and untimed mathematic skills. In addition the importance of language for successful number line estimation was studied. Thus, the progression in accuracy, format specificity and the relationship with timed and untimed mathematic achievement was investigated with a number line task on a 0-100 scale.

The first aim was to examine if number line estimation and development.

Firstly, the results of the analyses were in concordance with hypothesis 1 and in line with previous cross-sectional research (e.g., Ashcraft & Moore, 2012; Geary et al., 2008; Moeller et al., 2009; Muldoon et al., 2013; Sasanguie et al., 2013) revealing that and number line estimation errors on the 0-100 scale declined with age and instruction. There was a steady decrease in absolute errors from kindergarten to grade1 and a moderate decrease in errors in grade 2.

Secondly, there was mixed evidence for the format-independency of estimation (hypothesis 2). In line with the format-independency and studies of Barth and colleagues (2003), estimations became more accurate on all estimation tasks (using Arabic numerals, number words or dots as formats). However, in constrast with the format-independency, our findings suggested that more children had a linear distribution that was situated at earlier age than it was for number words. The number words in particular became format important in grade 2.

Thirdly, the shift from a logarithmic to a linear representation of numbers (hypothesis 3) was confirmed on 'group level'. Kindergarteners, the children in grade 1 and those at the beginning of grade 2 had a significant logarithmic representation of numbers on a 0-100 scale. In addition, there was a linear distribution for Arabic numbers, number words and dots in the middle of grade 2. These results are in line with the majority of Siegler and Booth (2004)'s findings on second graders, responding logarithmically on their 0-1000 number line. Additional analyses on an 'individual level' demonstrated that the linear representation of number words on a 0-100 number line occurred less often (in fewer children) and later (later period) than the representation of Arabic numbers in the same children.

Fiftly, the relationship between estimation and untimed and times math learning was studied (hypothesis 4). Especially untimed math learning could be predicted. These findings are in line with Sasanguie et al. (2013) that estimation on a number line is especially correlated with calculation skills and less with timed arithmetic or fact retrieval skills. Moreover, our findings revealed that number words became important in the prediction in grade 2. This might perhaps be explained by the inversion principle of two-digit number word names in Dutch (e.g., “een-en-zeventig”, literally “one-and-seventy”, for 71). Seron and Fayol (1994) showed that due to irregularities in the number word system, second graders from France made more errors on items comprising these numbers in different tasks (e.g., transcoding numbers from verbal to Arabic notation, transcoding numbers from verbal notation to representation with tokens, grammaticality judgements) compared to second graders from Wallonia. Dowker, Bala and Lloyd (2008) showed that Welsh speaking children (with a regular number word system) were better in magnitude comparison of two-digit numbers, but not in arithmetic compared to English speaking peers (with an irregular number word system) (Dowker, Bala, & Lloyd, 2008).

The second aim of this study, was to investigate whether language could explain some of the interindividual variation in the growth curves (hypothesis 5). Our analysis revealed that intelligence (assessed in kindergarten) explained a part of the variability in intercepts and slopes, whereas language (also assessed in kindergarten) explained variation when children enter the school system (in kindergarten) but not in the evolution of growth curves. These findings are in line with Ansari and colleagues (2003), meaning that language influenced the starting point but not the development or evolution of the estimation accuracy.

Even longitudinal studies have their limitations. Firstly, one of the limitations is that we relied on a limited set of numbers to be estimated in different formats. Several studies (Berteletti et al, 2010; Booth & Siegler, 2006) have used twice the amount of trials. In this study the function fits were calculated on 30 data points. Our decision to use three exercise trials and 30 test trials with 2, 3, 4, 6, 18, 25, 42, 67, 71 and 86 as stimuli was theoretically motivated since we were especially interested in the three formats. Nevertheless we have to mention the study by Gunderson, Ramirez, Beilock and Levine (2012) and Ebersbach et al. (2008), where even lesser numbers were used without resulting in instable responses of the children. Secondly, only half of the children were followed up in grade 2. This choice of following up only half of the children in grade 2 was motivated due to constraints in access to schools and the children attending them. The children did not drop out, so the missing data were random. Finally we have to acknowledge that not all authors consider the number line task as a task that measures numerical representations. Some of them (e.g., Barth & Paladino, 2011) see it as a measure of proportional judgment. In addition, evidence for a segmented linear model has been revealed by Ebersbach and colleagues (2008) and a M-shaped pattern was described by Ashcraft and Moore (2012) beginning in third graders' errors and fourth graders' latencies, suggesting that estimation comes to rely on a midpoint strategy, based on children's growing number knowledge (i.e., knowledge that 50 is half of 100). We did not run analyses in terms of shape of the distribution of estimations (bi-linear, M-shaped etc.) because of the limited estimation points in our task and because this was beyond the scope of this study.

Nevertheless, the current study has educational merits providing longitudinal evidence for the importance of familiarity with numbers leading to better estimation and untimed math proficiency. Perhaps, in line with Obersteiner, Reiss and Ufer (2013) a preventive support or larger amount of 'additional focusing on the position of numbers' for low performing kindergarteners can enhance their math skills. This needs to be addressed in future studies. The results also suggest that children enter kindergarten with different language skills, but language does not explain growth of estimation skills. Such knowledge is necessary in order to inform researchers and professionals about the value of testing language in kindergarten.

References

Ansari, D., Donlan, C., Thomas, M. S. C., Ewing, S. A., Peen, T., & Karmiloff-Smith, A. (2003). What makes counting count? Verbal and visuo-spatial contributions to typical and atypical counting development. *Journal of Experimental Child Psychology, 85,* 50-62. http://dx.doi.org/10.1016/S0022-0965(03)00026-2.

Ashcraft, M. H., & Moore, A. M. (2012). Cognitive processes of numerical estimation in children. *Journal of Experimental Child Psychology, 111*, 246–267. http://dx.doi.org/10.1016/j.jecp.2011.08.005.

Barner, D., Chow. K., & Yang, S. (2009). Finding one's meaning: A test of the relation between quantifiers and integers in language development. *Cognitive Psychology*, *58, 195-219*. http://dx.doi.org/10.1016/j.cogpsych.2008.07.001.

Barth, H., Kanwisher, N., & Spelke, E. (2003). The construction of large number representation in adults. *Cognition, 86,* 201-221. http://dx. doi.org/10.1037/a0013046.

Barth, H., & Paladino, A.M. (2011). The development of numerical estimation: Evidence against a representational shift. *Developmental Science 14, 125-135*. http://dx. doi.org/10.1111/j.1467-7687.2010.00962.x.

Baudonck, M., Debusschere, A., Dewulf, B., Samyn, F., Vercaemst, V., & Desoete, A. (2006). Kortrijkse Rekentest Revisie (KRT-R) [Kortrijk Arithmetic Test Revision]. Kortrijk: Revalidatiecentrum Overleie.

Bertelli, I., Lucangeli, D., Piazza, M., Dehaene, S., & Zorzi, M. (2010). Numerical estimation in preschoolers. *Developmental Psychology, 46*, 545-551. http://dx.doi.org/10.1037/a0017887.

Berteletti, I., Lucangeli, D., & Zorzi, M. (2012). Representation of numerical and non-numerical order in children. *Cognition, 124*, 304-313. http://dx.doi.org/101016/j.cognition.2012.05.015.

Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology*, *42*, 189-201. http://dx.doi.org/10.1037/0012-1649.41.6.189.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, *44,* 1–42. http://dx.doi.org/ 10.1016/0010-0277(92)90049-n.

Dehaene, S. (1997). *The Number Sense. How the Mind Creates Mathematics*. New York: Oxford University Press.

Dehaene, S. (2001). Précis of "the number sense". *Mind and Language, 16,* 16–32. http://dx.doi.org/ 10.1111/1468-0017.00154.

Dehaene, S., & Cohen, L. (1995). Towards an anatomical and functional model of number processing. *Mathematical Cognition, 1*, 83-120

Dehaene, S., Izard, V., & Piazza, M. (2005). Control over non-numerical parameters in numerosity experiments. Unpublished manuscript (available at http://www.unicog.org/…/DocumentationDotsGeneration.doc).

De Smedt, B., Noël, M-P., Gilmore, C., & Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends in Neuroscience and Education, 2,* 48-55. http://dx.doi.org/10.1016/j.tine.2013.06.001.

Desoete, A., & Roeyers, H. (2006). *Cognitieve Deelvaardigheden Rekenen (CDR)* [Cognitive Skills of Artithmetc]. Herenthals: VVL.

De Vos, T. (1992). Tempo Test Rekenen [Timed Fact Retrieval Test]. Nijmegen, The Netherlands: Berkhout.

Dowker, A., Bala, S., & Lloyd, D. (2008). Linguistic Influences on Mathematical Development: How important is the transparency of the counting system? *Philosophical Psychology, 21,* 523-538. http://dx.doi.org/10.1080/09515080802285511.

Ebersbach, M., Luwel, K., Frick, A., Onghena, P., & Verschaffel, L. (2008). The relationship between the shape of the mental number line and familiarity with numbers in 5 to 9 year old children: Evidence for a segmented linear model. *Journal of Experimental Child Psychology, 99,* 1-17. http://dx.doi.org/10.1016/jjecp.2007.08.006.

Geary, D. C., Hoard, M. K., Nugent, L., & Byrd-Craven, J. (2008). Development of number line representations in children with mathematical learning disability. *Developmental Neuropsychology, 33,* 277-299. http://dx.doi.org/10.1080/87565640801982361.

Grégoire, J., Noël, M., & Van Nieuwenhoven (2004). *Tedi-Math.* Antwerpen: Harcourt.

Gunderson, E.A., Ramirez, G., Beilock, S.L., & Levine, S.C. (2012). The relation between spatial skill and early number knowledge: the role of the linear number line. Developmental Psychology, 48, 1229-1241. http://dx.doi.org/10.1037/a0027433.

Hendriksen , J., & Hurks, P. (2009) *WPPSI-III-NL Wechsler Preschool and Primary Scale of Intelligence*. Pearson Assessment and Information B.V.

Holloway, I.D., & Ansari, D. (2009) Mapping numerical magnitudes onto symbols: The numerical distance effect and individual differences in children's math achievement. *Journal of Experimental Child Psychology, 103*, 17-29. http://dx.doi.org/10.1016/j.jecp.2008.04.001.

Huntley-Fenner, G., & Cannon, E. (2000). Preschoolers' magnitude comparisons are mediated by a preverbal analog mechanism. *Psychological Science, 11*, 147-152. http://dx.doi.org/ 10.1111/1467-9280.00230.

Kolkman, M.E., Kroesbergen, E.H., & Leseman, P.P.M. (2013). Early numerical development and the role of non-symbolic and symbolic skills. *Learning and Instruction, 25*, 95-103. http://dx.doi.org/10.1016/j.learninstruc.2012.001.

Kort, W., Schittekatte, M., & Compaan,E.,(2008) *CELF-4-NL Test voor diagnose en evaluatie van taalproblemen. Handleiding [Test for the evaluation of language problems. Manual]*. Amsterdam: Pearson.

Moeller, K., Pixner, S., Kaufmann, L., & Nuerk, H-C. (2009). Children's early mental number line: logarithmic of decomposed linear? *Journal of Experimental Child Psychology, 103,* 503-515. http://dx.doi.org/10.1016/j.jecp.2009.02.006.

Muldoon, K., Towse, J., Simms, V., Perra, O., & Menzies, V. (2013). A longitudinal analysis of estimation, counting skills and mathematical abilities across the first school year. *Developmental Psychology, 49,* 250-257. http://dx.doi.org/ 10.1037/a0028240.

Mundy, E., & Gilmore, C. K. (2009). Children's mapping between symbolic and non-symbolic representations of number. *Journal of Experimental Child Psychology, 103,* 490–502. https://dspace.lboro.ac.uk/2134/8765.

Negen, J., & Sarnecka, B.W. (2012). Number-Concept Acquisition and General Vocabulary Development, *Child Development*, 83, 2019-2027. http://dx.doi.org/10.1111/j.1467-8624.2012.01815.x.

Obersteiner, A., Reiss, K., & Ufer, S. (2013). How training on exact or approximate mental representations of number can enhance first-grade students' basic number processing and arithmetic skills. Learning and Instruction, 23, 125-135. http://dx.doi.org/10.1016/j.learninstruc.2012.08.004.

Praet, M., Titeca, D., Ceulemans, A., & Desoete, A. (2013). Language in the prediction of arithmetic in kindergarten and grade 1. *Learning and Individual Differences*, http://dx.doi.org/10.1016/j.linidif.2013.07.003.

Purpura, D.J., Hume, L.E., Sims, D,C., & Lonigan, C.J. (2011). Early literacy and early numeracy: The value of including early literacy skills in the prediction of numeracy development. *Journal of Experimental Child Psychology, 110,* 647-658, http://dx.doi.org/10.1016/j.jecp.2011.07.004.

Romano, E., Babchishin, L. Pagani, L.S., & Kohen, D. (2010). School readiness and later achievement: Replication and extension using a nationwide Canadian survey. *Developmental Psychology, 46,* 995-1007. http://dx.doi.org/10.1037/a0018880.

Sarnecka, B.W., Kamenskaya, V.G., Yamana, Y., Ogura, T., & Ydovina, Y.B. (2007). From grammatical number to exact numbers: early meanings of one, two and three in English, Russian and Japanese, *Cognitive Psychology, 55,* 136-168. http://dx.doi.org/10.1016/j.cogpsych.2006.09.001.

Sasanguie, D., Göbel, S.M., Moll, K., Smets, K., & Reynvoet, B. (2013). Approximate number sense, symbolic number processing, or number-space mappings: What underlies mathematics achievement? *Journal of Experimental Child Psychology, 114,* 418-431. http://dx.doi.org/10.1016/j.jecp.2012.10.012.

Semel, E., Wiig, E. H., & Secord W.A.(2008). *CELF 4 Nl Clinical Evaluation of language Fundamentals*. Amsterdam :Pearson.

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development, 75,* 428-444. http://dx.doi.org/10/1111/j.1467-8624.2004.00684.x.

Siegler, R.S., & Opfer, J.E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science, 14,* 237-243. http://dx.doi.org/10.1111/1467-9280.02438.

Schneider, M., Heine, A., Thaler, V., Torbeyns, J., De Smedt, B., Verschaffel, L., . . . Stern, E. (2008). A validation of eye movements as a measure of elementary school children’s developing number sense. *Cognitive Development, 23,* 409-422. http://dx.doi.org/10.1016/j.cogdev.2008.07.002.

Seron, X., & Fayol, M. (1994). Number transcoding in children: A functional analysis. *British Journal of Developmental Psychology, 12,* 281-300. http://dx.doi.org/10.1111/j.2044-835X.1994.tb00635.x.

Slusser, E.B., Santiago, R.T., & Barth, H.C. (2013). Developmental change in numerical estimation. *Journal of Experimental Psychology, 142,* 193-208. http://dx.doi.org/10.1037/a0028560.

Stock, P., Desoete, A., & Roeyers, H. (2010). Detecting children with arithmetic disabilities from kindergarten: Evidence from a three year longitudinal study on the role of preparatory arithmetic abilities. *Journal of Learning Disabilities*, *43*, 250-268. http://dx.doi.org/10.1177/0022219409345011.

Wechsler, D., Kort,W., Schittekatte, M., Bosmans, M., Compaan, E.L., Dekker, P.H., &Verhaeghe, P. (2002). Wechsler Intelligence Scale for Children-III-Nl. Amsterdam, The Netherlands: Harcourt.

Wiese, H. (2003). Iconic and non-iconic stages in number development: the role of language. *Trends in Cognitive Science, 7,* 385-390. http://dx.doi.org/10.1016/S1364-6613(03)00192-X.

Table 1: Percentage of Absolute Error (PAE) for the different modalities

	Period 1 M *(SD)*	**Period_2** M *(SD)*	**Period_3** M *(SD)*	**Period_4** M *(SD)*	**Period_5** M *(SD)*
Dots	24.69	18.80	15.71	13.80	13.49
	(8.91)	(6.46)	(8.03)	(5.64)	(4.76)
Arabic numbers	24.18	18.64	12.48	10.47	9.58
	(8.93)	(6.69)	(6.91)	(6.21)	(5.52)
Number words	25.03	19.10	12.36	11.39	9.70
	(10.78)	(7.36)	(6.16)	(5.91)	(5.05)
Total PAE	24.92	18.90	13.57	11.89	10.95
	(8.74)	(5.99)	(6.28)	(5.22)	(4.50)

Table 2: Distribution of estimations from kindergarten till grade 2.

	R^2lin	*p*lin	R^2log	*p*log	*t*	*P*
Kindergarten						
Total test	.729	.002	**.959**	.000	*t*(129)=-12.712	.000*
Arabic Numbers	.751	.001	**.946**	.000	*t*(129)= -8.561	.000*
Number words	.766	.001	**.968**	.000	*t*(129)=5.935	.000*
Dots	.740	.001	**.927**	.000	*t*(129)= -8.304	.000*
Grade1 time 2						
Total test	.843	.000	**.970**	.000	*t*(129)=-8.096	.000*
Arabic Numbers	.848	.000	**.976**	.000	*t*(129)=-8.777	.000*
Number words	.894	.001	**.984**	.000	*t*(129)=-6.497	.000*
Dots	.737	.000	**.751**	.001	*t*(129)=-4.112	.000*
Grade1 time 3						
Total test	.899	.000	**.961**	.000	*t*(126)= -5.962	.000*
Arabic Numbers	.911	.000	**.969**	.000	*t*(126)=-5.456	.000*
Number words	.917	.000	**.976**	.000	*t*(126)=-5.595	.000*
Dots	.833	.000	**.854**	.000	*t*(126)=-3.070	.003*
Grade2 time 4						
Total test	.938	.000	**.956**	.000	*t*(58)=-2.102	.040*
Arabic Numbers	.965	.000	.933	.000	*t*(58)=-.034	.973
Number words	.952	.000	**.955**	.000	*t*(58)=-2.602	.012*
Dots	.914	.000	**.894**	.000	*t*(58)=-2.297	.025*
Grade2 time5						

Total test	.977	.000	.892	.000	$t(61)=-.997$	.323
Arabic Numbers	**.992**	.000	.888	.000	$t(61)=2.701$	.009*
Number words	**.992**	.000	.890	.000	$t(61)=2.953$	.004*
Dots	**.932**	.000	.863	.000	$t(61)=2.029$	.047*

* $p<.05$

Table 3: Percentage of children having a linear and logarithmic representation

	Lin representation	Log representation	No representation
Kindergarten			
Arabic	6.8%	49.2%	42.4%
Number words	11.4%	45.5%	41.7%
Dots	9.8%	50%	38.6%
Start of grade 1			
Arabic	20.2%	74.4%	5.4%
Number words	29%	65%	6%
Dots	27.9%	44.2%	27.9%
End of grade 1			
Arabic	32.2%	63.8%	3.9%
Number words	32.3%	61.4%	6.3%
Dots	35.4%	55.2%	9.4%
Start of grade 2			
Arabic	55.9%	42.4%	1.7%
Number words	37.3%	61%	1.7%
Dots	39%	54.2%	6.8%

Middle grade 2			
Arabic	62.3%	32.8%	5%
Number words	60.7%	36.0%	3.3%
Dots	59.0%	32.8%	8.2%

Table 4: Predictions with math learning as outcome in kindergarten, grade 1 and grade 2

	Unstandardized Coefficients	β	t	*p*
Kindergarten (T1)				
Constant	14.330		11.209	.000
PAE Dots	-.008	-.014	-0.126	.900
PAE Arabic Numbers	-.211	-.363	-2.517	.013
PAE Number words	-.056	-.117	-0.874	.384
Grade 1 T2 KRT-R				
Constant	50.030		21.149	.000
PAE Dots	-.281	-.221	-1.646	.102
PAE Arabic Numbers	-.234	-.167	-1.01	.315
PAE Number words	.057	.045	0.276	.783
Grade 1 T3 TTR				
Constant	25.947		17.899	.000
PAE Dots	-.052	-.066	-0.534	.595

PAE Arabic Numbers	-.131	-.142	-0.966	.336
PAE Number words	-.061	-.058	-0.382	.703
Grade 1 T3 CDR				
Constant	31.761		24.419	.000
PAE Dots	.003	.004	0.032	.975
PAE Arabic Numbers	-.063	-.069	-0.517	.606
PAE Number words	-.414	-.406	-2.906	.004
Grade 2 T5 KRT-R				
Constant	35.536		9.817	.000
PAE Dots	-.153	-.072	-.543	.590
PAE Arabic Numbers	-.414	-.226	-1.315	.194
PAE Number words	-.616	-.307	-2.016	.048
Grade 2 T5 TTR			10.248	
Constant	59.332			
PAE Dots	.207	.072	.462	.646
PAE Arabic Numbers	.075	.030	.152	.880

PAE Number words	-1.100		-.399	-2.255	.028

*p ≤ .05 Note. PAE = Percentage Absolute Error

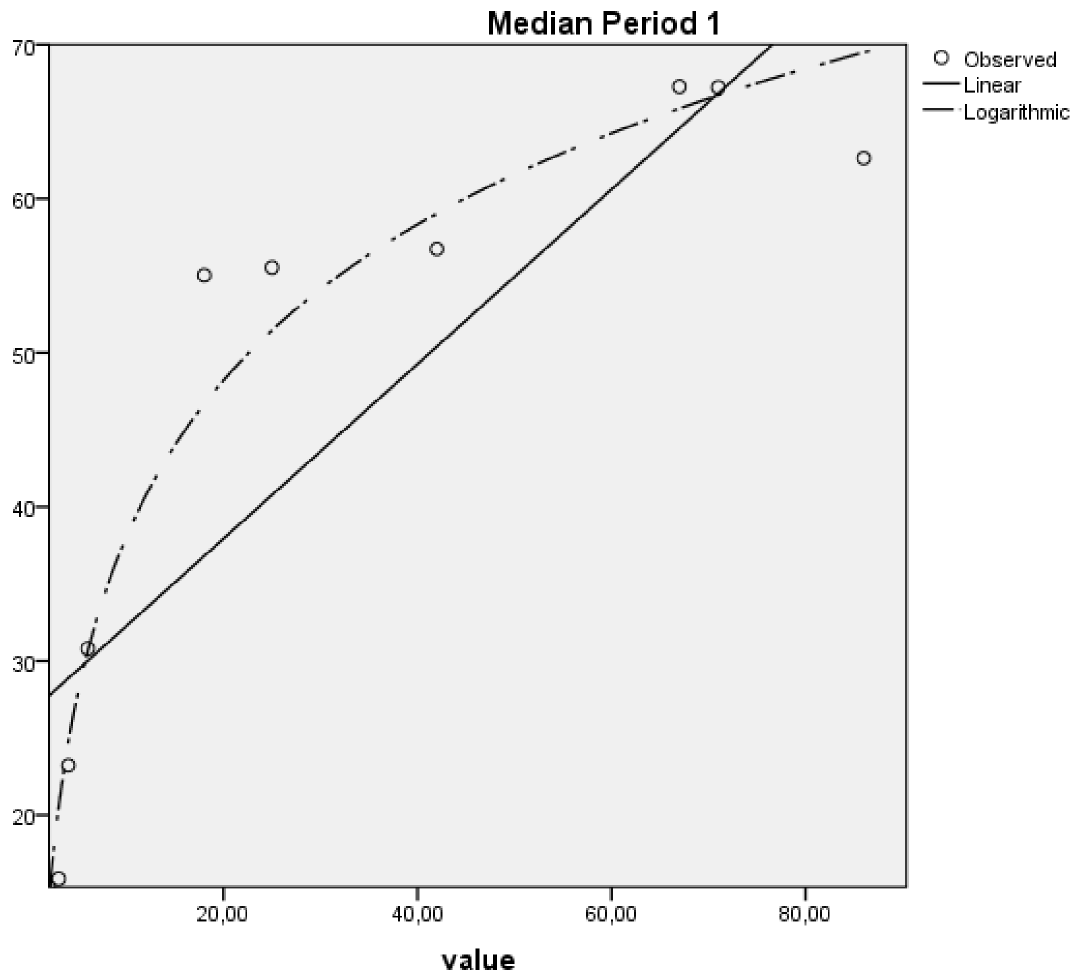

Figure 1. Representation in kindergarten

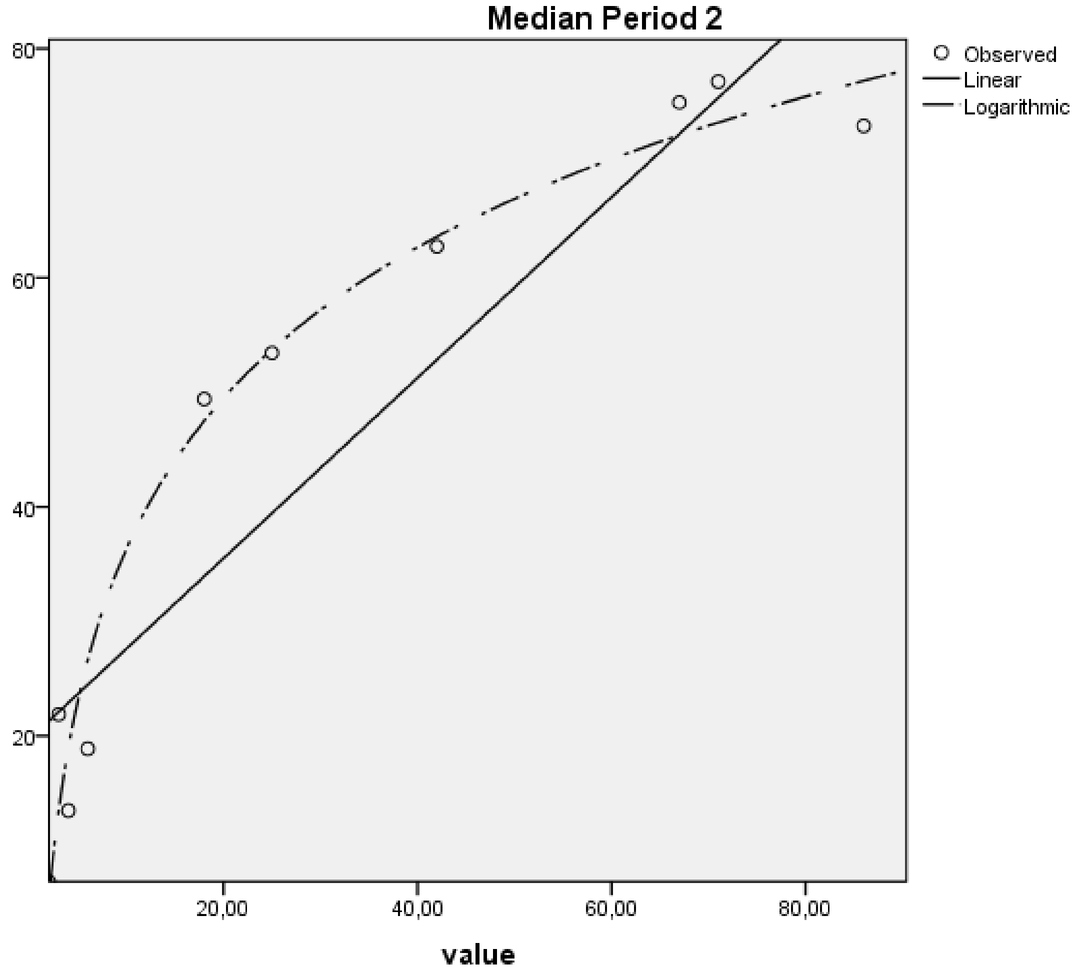

Figure 2. Representation at the beginning of grade 1

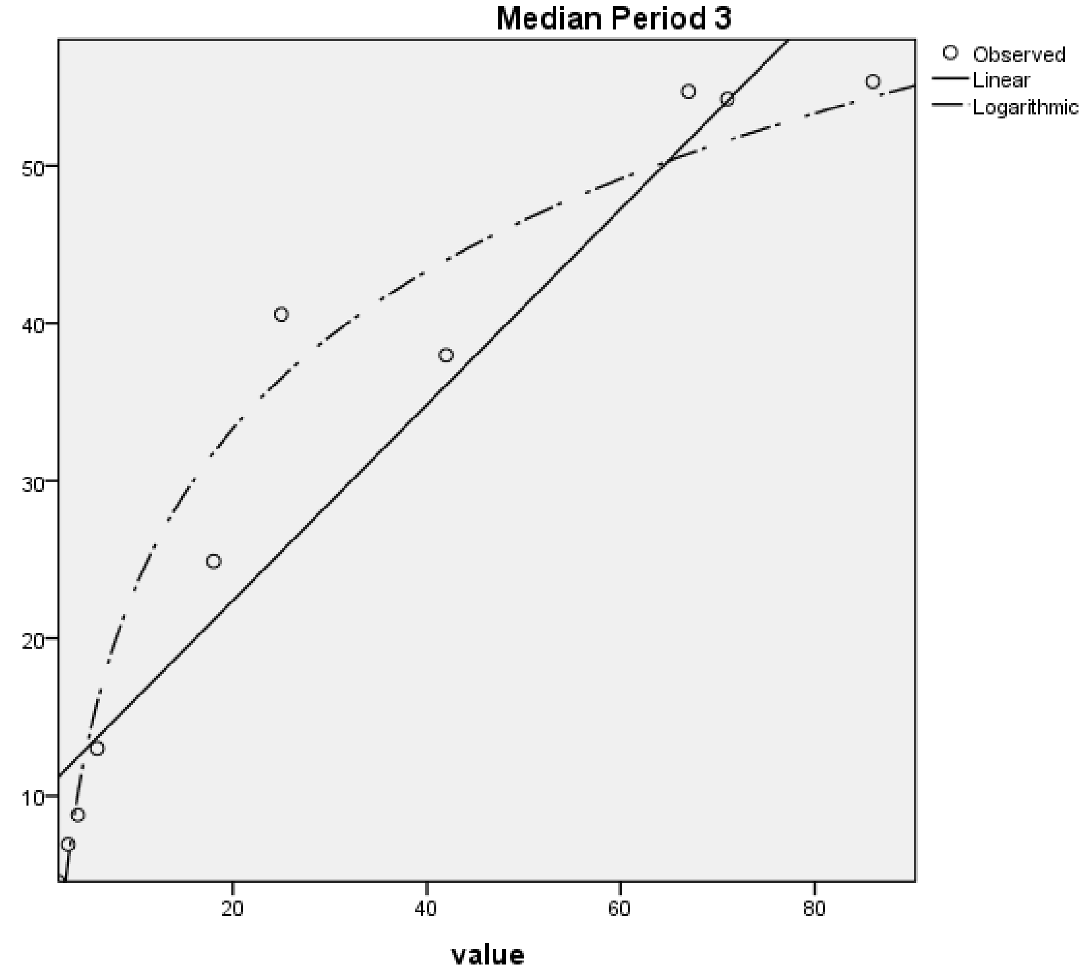

Figure 3. Representation at the end of grade 1

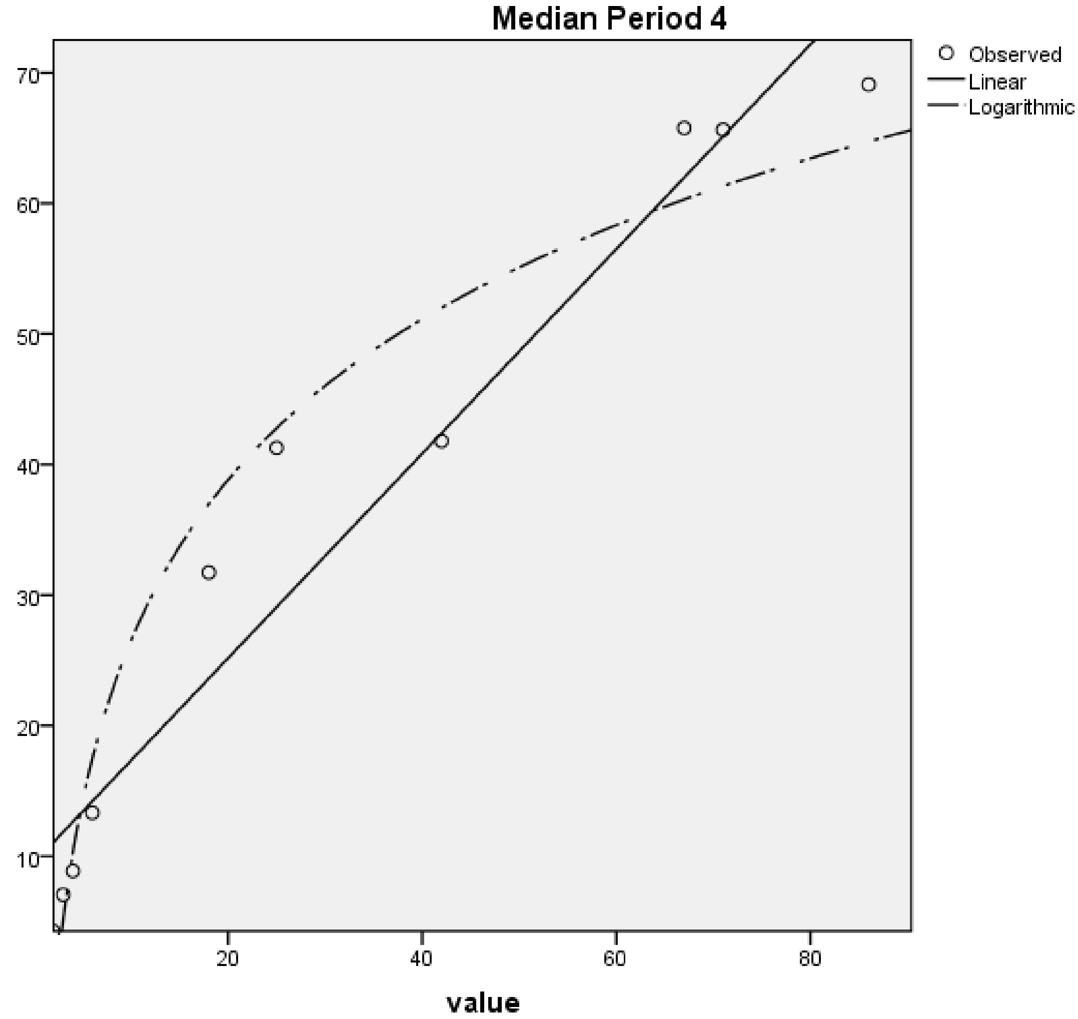

Figure 4. Representation at the start of grade 2

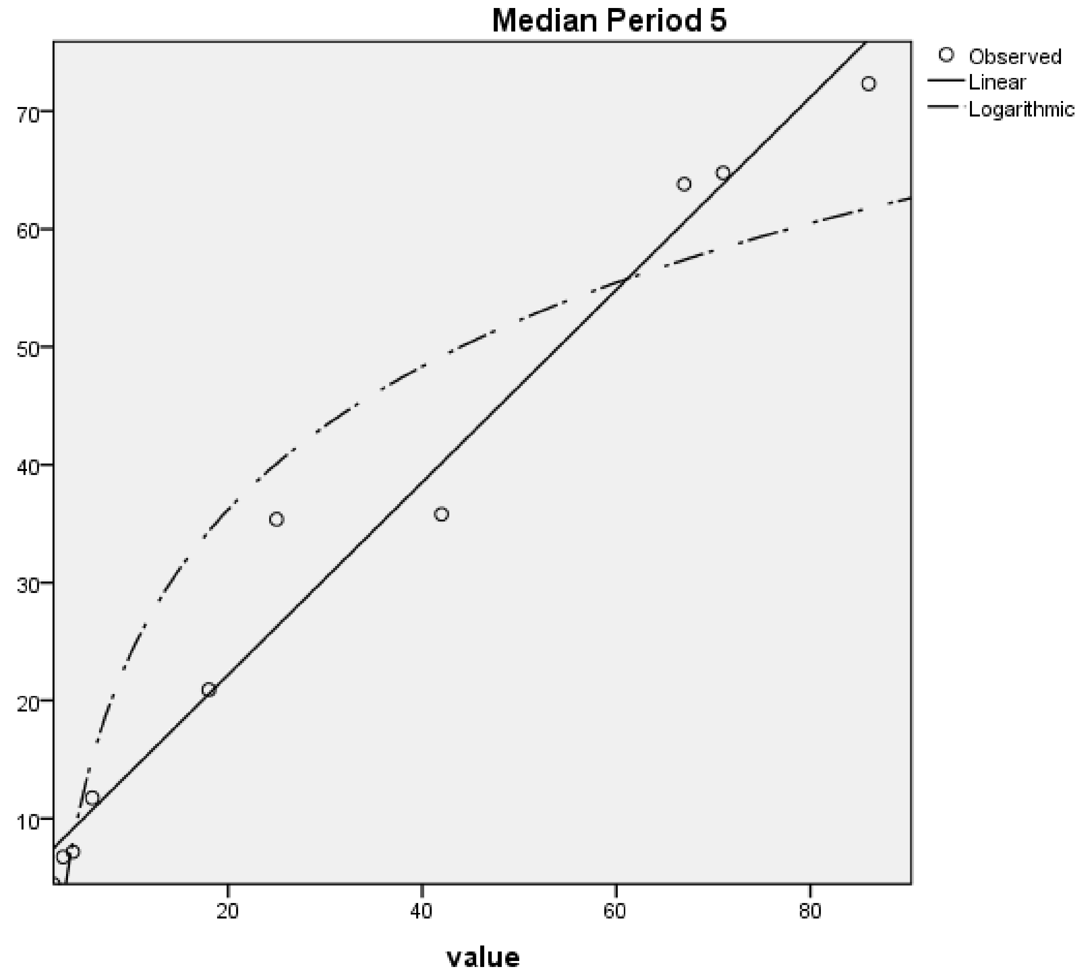

Figure 5. Representation in the middle of grade 2

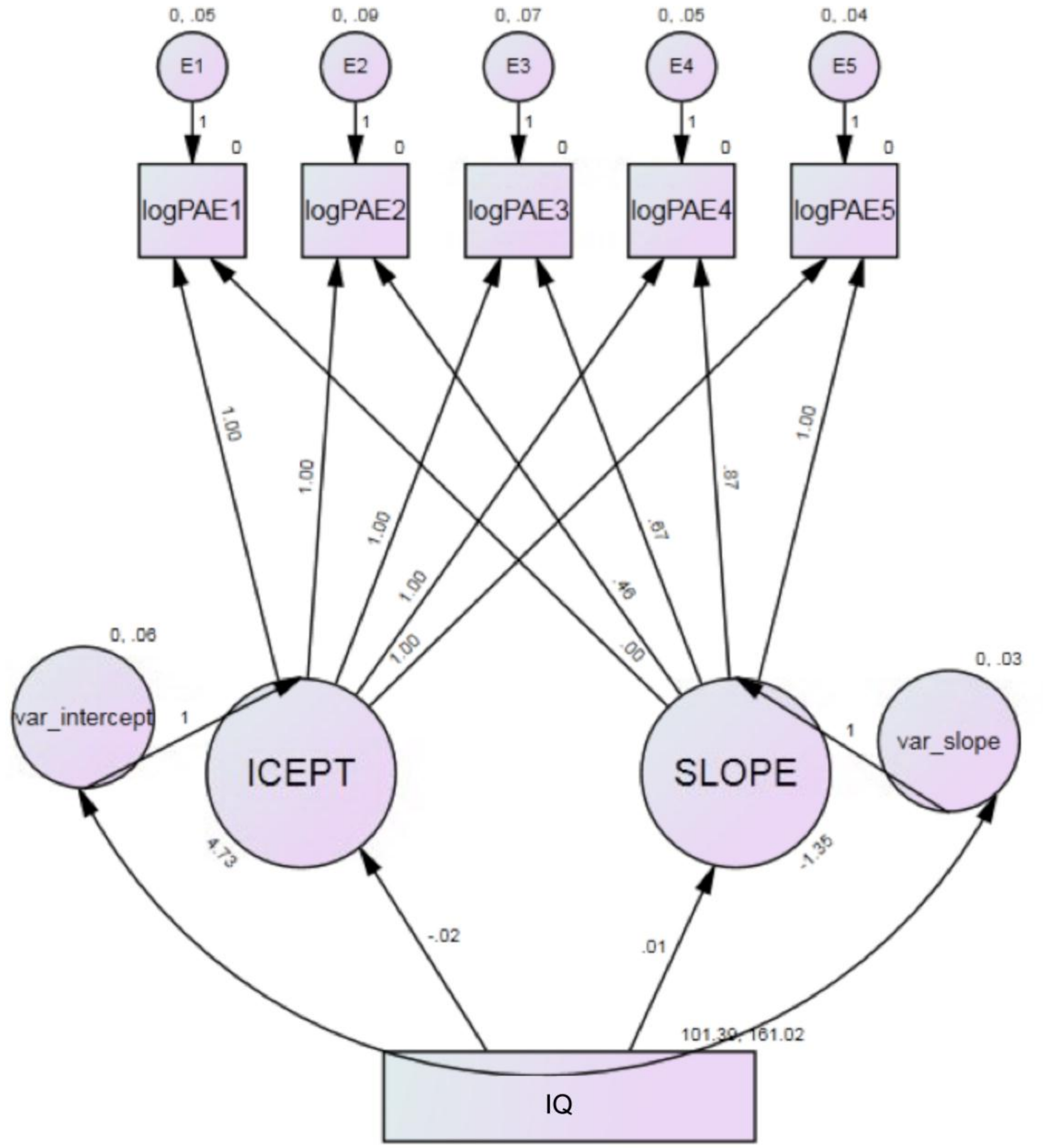

Figure 6. Latent Growth curve on estimation accuracy with intelligence as covariate

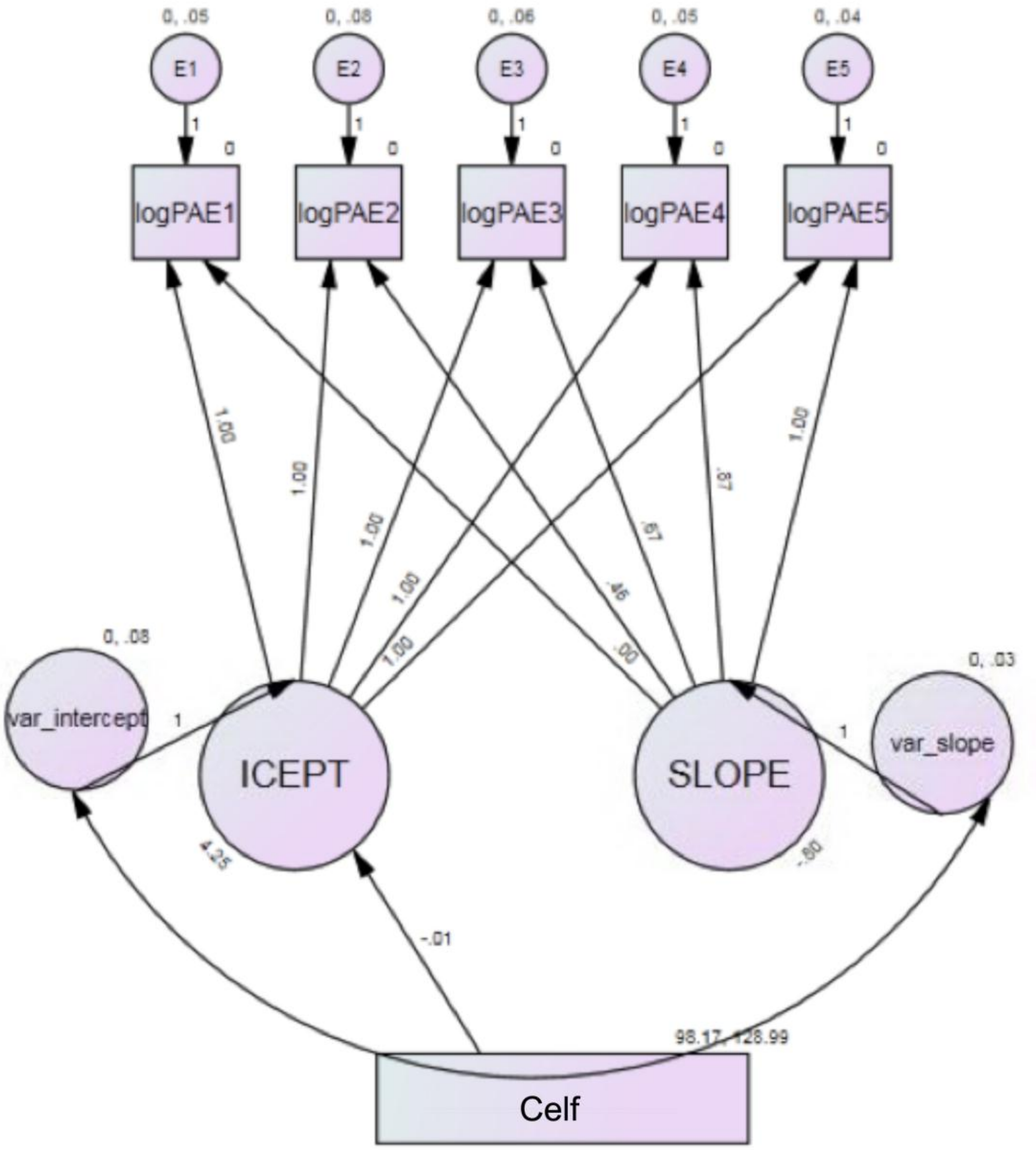

Figure 7. Latent Growth curve on estimation accuracy with language as covariate