EDUCATIONAL ASSESSMENT & EVALUATION | RESEARCH ARTICLE

The role of PS ability and RC skill in predicting growth trajectories of mathematics achievement

Alvin Vista1,*

Abstract: There are relatively few studies in Australia and South-East Asian region that combine investigating models of math growth trajectories with predictors such as reasoning ability and reading comprehension skills. Math achievement is one of the major components of overall academic achievement and it is important to determine what factors (especially domain-general factors) predict levels of achievement over time. This study presents large-scale data (N = 5,886) from Australia to examine the trajectories of growth in math achievement and how problem solving (PS) ability and reading comprehension (RC) skills predict this growth among government school students (grades 3–8) in Victoria. Latent growth modelling showed that PS ability predicts growth in math achievement and this relationship is partially mediated by RC skill. Both PS ability and RC skill predict initial status in math achievement, but only PS ability predicts growth. The data also fit a model in which an improvement in general reasoning ability of students allows those with lower initial levels in math achievement to catch up. This can be interpreted as evidence that improving growth rates in PS ability may lift growth rates in math achievement.

Subjects: Educational Psychology; Educational Research; Research Methods in Education

Keywords: assessment; item response theory; large-scale assessment; Rasch model; reading comprehension; problem solving; growth models

ABOUT THE AUTHOR

Alvin Vista was professionally engaged as part of a research team at the Assessment Research Centre, University of Melbourne in conducting the study for this paper, and where he has completed his PhD in Educational Measurement. Alvin is currently a research fellow at the Australian Council for Educational Research. His main research interests are in test development and educational measurement in cross-cultural settings.

The data for this study were collected as part of a broader research project called the Assessment and Learning Partnerships (ALP) conducted by the Assessment Research Centre. The collaboration for the data collection process and the development of the research instruments included the contributions of M. Pavlovic, N. Awwal and P. Robertson as the research team under the supervision of Patrick Griffin and Esther Care as the principal investigators.

PUBLIC INTEREST STATEMENT

This paper looks at the relationship between maths achievement, reading comprehension (RC) skill and problem solving (PS) ability. An increasing linear trend has been established for maths achievement. There is statistically significant variation in both the initial maths achievement levels and growth rates in maths achievement. RC skill affects the initial status in maths achievement more than it affects the trajectory of growth. The rate of growth in PS ability significantly predicts the rate of growth in maths achievement. However, PS ability is less affected by time-related factors compared with growth in maths achievement.

© 2016 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.
1. Background
Measuring growth in achievement is central to educational measurement. Single measures may be useful when the goal is to simply assess at which level in any particular area of study a student stands. But in order to assess improvement and rate of learning, we need to have measures of growth that allow us to determine not just the current level but also the rate of growth so that we can project the level of achievement over time. In conjunction with measuring growth, it is also important to determine if certain factors predict this growth. Mathematics achievement is one of the major components of overall academic achievement (Duncan et al., 2007) and predicts real-life attainment beyond what is attributable to general intelligence (Rivera-Batiz, 1995). It is therefore not surprising that a substantial number of studies on academic achievement focus on math as one of its most important domains.

There have been Australian studies on the links between math achievement and demographic factors (e.g. Clarkson (2006), Clarkson and Leder (1984), Hemmings, Grootenboer, and Kay (2011)), as well as the link between these factors and more general cognitive abilities (e.g. Carstairs, Myors, Shores, and Fogarty (2006), Dandy and Nettelbeck (2002)), but most of these studies are based on limited samples. More importantly, there is only a handful of Australian studies that look at changes in math achievement over time. Two notable studies with large-scale national data are by Marks and Ainley (1997), who looked at Australian literacy and numeracy data among grades 7 to 9 students over 20 years, and by Rothman (2002), who extended the study by including younger students and increasing the time span to almost 25 years. Both of these studies looked at the effects of demographic variables (including language background) on achievement trends. However, they did not examine the trajectory of math growth in a way that would enable theoretical models to be developed nor did they look at what could predict the trajectory of such growth. There has been no Australian study of comparable scale that has focussed on math achievement growth trajectory, and in particular, no study on the association between this growth trajectory and general cognitive abilities. Even internationally, there are relatively few studies that combine investigating models of math growth trajectories with predictors such as reasoning ability and reading comprehension (RC) skills (Lee, 2010).

This gap in research is what this study attempts to fill. It is well established that problem solving (PS) and general cognitive abilities strongly predict math achievement (Deary, Strand, Smith, & Fernandes, 2007; Geary, 1994; Jensen, 1998; Taub, Keith, Floyd, & McGrew, 2008). This study, therefore, extends what is known about this association between general cognitive abilities and math achievement by investigating the dynamics of how reasoning ability affects the trajectory of growth.

The inclusion of RC skill as predictor in models of growth trajectories bridges the gap in research by providing factors that can explain the mechanisms of growth. Studies on growth trajectories for math, reading or both have been reported (Guglielmi, 2012; Lee, 2010). However, as in the case of Lee’s study, the models are useful in formalising the mechanisms of math and reading growth over time but they do not explore the association between these two constructs, nor do they look at possible external factors that can affect the mechanisms of growth. Primi, Ferrão, and Almeida (2010) address this need to include predictors that can explain growth trajectories, but their study is still limited in that no covariates were included in their models. Guglielmi (2012) did explore the association between math and reading achievement while also including relevant covariates that can inform on the association between the constructs of interest, but it is focussed on English-language learners. This paper aims to combine the general research thrust of modelling the trajectory of math growth of a more general and nationally representative sample, in a framework that includes both direct predictors and possible covariates based on naturally occurring groups.

2. Reading comprehension skill and test language
This study capitalises on the rare opportunity to have a large and representative assessment sample data from government schools that include measures of PS ability. Nationwide tests have numeracy and literacy components, but the inclusion of a reasoning ability component is relatively uncommon.
This is important because the inclusion of this reasoning component in this study allows measurement of a construct that affects both numeracy and literacy components of academic performance while being relatively independent of schooling.

A study by Parrila, Aunola, Leskinen, Nurmi, and Kirby (2005) elucidated the mechanisms of growth for reading development, but they did not extend their models to include other domains of learning (such as math) or domain-general cognitive processes. There is limited understanding of what factors might influence how math achievement gaps narrow and to what extent these factors influence such rate. Exploring the mechanism of these relationships could provide a better understanding of how growth in math achievement is affected by important cognitive processes such as reasoning and language skills. In turn, a better understanding of these mechanisms could inform targeted interventions in areas that result in optimal growth.

It is important to investigate the link between language proficiency and performance in math yet research on the association between language proficiency and trends in math achievement, is rare. A comprehensive review on trends in math achievement in the US by Tate (1997), spanning some 15 years, found only one study that looked at language proficiency as predictor. There is similar lack of studies on the predictors of math achievement in Australia, especially studies that look at predictors of math achievement growth (Hemmings et al., 2011).

The measurement of RC skill is part of literacy assessment, which involves both reading and writing as the two main components of literacy. From an assessment perspective, literacy is defined in the curriculum as consisting of two main processes: comprehension and composition of texts (Australian Curriculum, Assessment & Reporting Authority, 2012). RC is the component and more delimited construct that is used in this study. As a construct, RC can be defined as a “conceptualisation of skills and knowledge that comprise the ability to make meaning of text” (Morsy, Kieffer, & Snow, 2010, p. 3).

In order to develop models of growth trajectories that take into account Australia’s classroom diversity, this study includes an analysis of how the dynamics of this association is influenced by proficiency on the test language. Important language skills (e.g. decoding and comprehension, phonological processing) either directly predict math performance (Fuchs et al., 2006; Seethaler & Fuchs, 2006) or at least covary with it (Hart, Petrill, Thompson, & Plomin, 2009; Vilenius-Tuohimaa, Aunola, & Nurmi, 2008), thus the test language has direct implications for non-native speakers of this language. This effect of test language on math performance of non-native speakers is not uniform because of variation in language loadings for different types of math skills being measured (for example, arithmetic calculation may have less language loads compared with more complex word problems) (see, for example, findings in Fuchs et al. (2005, 2006)). Regardless of differential effects that depend on skills being measured and type of test items, it can be argued that skills in the test language would have some influence on the performance in these tests, just as reading skills overlap with general cognitive ability in its loading on math ability (Hart et al., 2009).

3. General reasoning ability and mathematics ability

General reasoning ability is broadly defined and involves both inductive and deductive reasoning, as well as divergent and convergent thinking skills. While these skills are essential in curricular domains such as math, tests of general reasoning ability are usually constructed to be domain-independent and generally do not require specific math knowledge. In this study, the PS test used measures general reasoning ability and creative thinking skills; the test does not require any specialised math skills. However, it is reasonable to expect that performance on a general PS test and a math ability test will be highly correlated. This association between general reasoning ability and math performance is influenced by a number of factors. For example, performance on measures that test visual-spatial working memory correlate with math achievement but the effect is mediated by fluid intelligence (Kyttälä & Lehto, 2008). Fluid intelligence (Gf) is one of two general factors of intelligence put forward more than 60 years ago by Cattell (1941, 1963), the other being crystallised intelligence.
intelligence (Gc). In what is now known as the GF-Gc Theory (Carroll, 1984; Cattell, 1963), the two factors differ mainly on how they depend on specific knowledge. GC is entangled with school-based skills and areas of learning, making the measurement of GC difficult to separate from the measurement of achievement (Cattell, 1963), while GF is conventionally accepted to be domain-general even if it strongly correlates with specific school-based skills (e.g. mathematical reasoning, Floyd, Evans, and McGrew (2003)). GF and working memory predict multi-tasking performance (Konig, Buhner, & Murling, 2005), which is crucial to math achievements.

This study, however, is not focused solely on GF as a predictor of math achievement. By using a measure that assesses general reasoning ability, this study takes into account both fluid and GC. Consequently, the measures of reasoning ability to be used in this study are not strictly nonverbal and in fact a considerable proportion carries significant verbal loads. This dual and inclusive approach may offer unique benefits compared with using language-free measures of GF in examining mediating effects of language proficiency. An Australian study that examined the effect of linguistic background on performance in an intelligence test found that language could be just one of two independent factors that have an effect on cognitive abilities and that English proficiency operates independent of sociocultural factors (Carstairs et al., 2006). In their study, group differences exist between English-speaking background (ESB) and non-English-speaking background (NESB) participants even if the NESB participants speak English as their first language. This shows that language proficiency can influence the performance on language-dependent portions of an intelligence test while sociocultural factors also exert an influence on the nonverbal portions (Carstairs et al., 2006).

3.1. Measuring math achievement

Benchmarks for measuring math achievement in the levels of schooling that are within the context of this study are based on standardised and national measures of math both in Australia and across the world. There is much debate on how math achievement should be defined, what measures are appropriate and how levels of achievement should be delineated. However, these wide range of issues are beyond the scope of this study and as such, discussions will be limited to what is the current de facto standard of measuring math achievement and how these measures are used within the context of each country’s governing institutions to define the levels of achievement. In other words, this study does not seek to redefine the measurement and structure of the construct that is math achievement, but is instead focussed on the actual measurement data. This means that the study utilises test data from well-established, preferably large-scale and standardised measures of math achievement.

In Australia, national testing is comparatively recent and the main measure of math ability is the numeracy test within a much larger system of national testing called the National Assessment Program—Literacy and Numeracy (NAPLAN). NAPLAN is the main instrument for Australia’s assessment programme that evaluates educational outcomes at the national level (Ministerial Council for Education, Early Childhood Development & Youth Affairs, 2009). Beginning in 2008, NAPLAN is administered annually for students in years 3, 5, 7 and 9 in four domains of learning: reading, writing, language conventions (spelling, grammar and punctuation) and numeracy (Australian Curriculum Assessment & Reporting Authority, 2011).

Although the available data and scope of population being tested in NAPLAN are comprehensive, there is an important limitation that needs to be considered in using it as a main source of data for this study. As mentioned earlier, in any given year, NAPLAN is only administered to years 3, 5, 7 and 9. This is a major logistical limitation and thus, for this study, a parallel and concurrently validated measure of math achievement was used—the Assessment and Learning Partnerships (ALP) tests (for details, see Assessment Research Centre, (2004, 2011)).

3.2. Measuring general reasoning ability

There is a need to delineate the measurement of general reasoning ability in the curricular and school-level contexts by framing it as a measure of PS ability because reasoning ability is an abstract
and relatively broad construct. By framing the test as a measure of PS ability, we limit the scope of the construct that we need to measure and are able to take advantage of an existing framework that specifies the curriculum standards.

The instrument for measuring PS ability in this study is a multiple choice test but it maximises the information about reasoning processes by applying findings from modern test development. In mathematical PS, for example, a task that is constructivist in approach may allow or even encourage the students to employ their own strategies based on their collection of math concepts, in contrast to tasks that require a specific mathematical approach to be demonstrated (Clarke, 1992). The PS items in the ALP tests used for this study were designed such that they are open-ended in terms of strategies needed to solve them. In other words, they do not require fixed concepts or rigid item-specific processes to solve.

Another important issue is that RC skill should not interact with PS ability. As the main predictor in the study design, the measure of PS ability (and thus the scope-limited measure of reasoning skills) must not be systematically biased with respect to RC skill. To address this issue, the instrument needs to have English reading loads that are not detrimental to the performance in this instrument. A purely language-free instrument would be ideal, but it would not be feasible because the instrument has to fit within the Victorian curricular specifications. Nevertheless, the study design incorporates statistical methods to check on whether the instrument exhibits systematic bias for any of the groups. This is done through a differential item functioning analysis on the PS instrument and by designing the instrument such that it does not require any specialised math skills that are specific to the grade level of students.

4. Research questions

Given the framework that math learning is linked with general reasoning ability (e.g. Casey, Pezaris, and Nuttall (1992), Hart et al. (2009), Primi et al. (2010)), this paper extends this by seeking to establish that the mechanism of this association between PS ability and math achievement is not uniform longitudinally. Further, this paper also aims to define the role of RC skill in this framework. Therefore, the two major research questions are: “What is the rate of growth in math achievement level as measured at three timepoints?” and “Do PS ability and RC skill predict the initial status and rate of growth in math achievement?” with sub-questions on the type of growth (linear or not) and the dynamics of growth (slowing or increasing over time). Results from a comparable study, but one that focuses on reading development (Parrila et al., 2005), showed that growth can be nonlinear. The number of timepoints in this study is limited but a linear model can still be explicitly tested for significant deviations from linear trajectories.

Findings from the literature show a positive association between initial level and subsequent level of math achievement (Jordan, Hanich, & Kaplan, 2003; Jordan, Kaplan, Ramineni, & Locuniak, 2009; Lerkkanen, Rasku-Puttonen, Aunola, & Nurmi, 2005) and we predict that the same association will be supported by results from this study. Positive correlations between reasoning ability and math (Fuchs et al., 2010; Geary, 1994, 2011), reasoning ability and RC (Evans, Floyd, McGrew, & Leforgee, 2002; Riding & Powell, 1987; Vellutino, 2001), and RC and math (Jordon, Kaplan, & Hanich, 2002; Lerkkanen et al., 2005) are well reported in the literature. We expect these positive associations to be supported by our results, but we also aim to put these three constructs together and elucidate the relationships among them.

The answers to the research questions in this study can have important implications for educational interventions that depend on a better understanding of how growth in maths achievement is affected by important cognitive processes such as reasoning and language skills.

4.1. Main design for modelling latent growth

The longitudinal study design and the way the data were collected have features that are suitable for an analysis of change to be conducted (see Singer and Willett (2003)): (a) multiple waves of data
(at least three to enable testing for linear growth), (b) a time-structured data-set with uniform and equally spaced timepoints and (c) continuous interval-scale outcome measures.

Among the more sophisticated ways of analysing longitudinal data is through linear growth modelling (LGM) with latent variables. This LGM approach can be traced back to seminal works by Rao (1958), Tucker (1958) and, under an SEM perspective as implemented in this study, works by Bollen and Curran (2006), McArdle (1987, 2009), McArdle and Epstein (1987), Meredith and Tisak (1990), and Muthén (1991). The two main advantages of using LGM under an SEM approach over OLS regression to analyse longitudinal data include more flexibility on the treatment of measurement error variances, and the capability to analyse change over time for individuals rather than only considering group means (Kline, 1998).

The simplest longitudinal growth model is that of linear change over time, based on three or more observations of the sample cases across time periods. This model is described in detail elsewhere (e.g. Bollen and Curran (2006), Duncan, Duncan, Strycker, Li, and Apert (2006), Kline (1998), Muthén (1991), Singer and Willett (2003)) but is summarised below in the manner in which it was implemented in this paper.

Suppose we want to model a linear growth for the same individual \( i = 1–N \) by measuring over a certain time period (\( t = 1–3 \) for this study). The unconditional model trajectory equation in scalar expression is:

\[
y_k = \eta_i + \eta_{\beta} \lambda_k + \varepsilon_k
\]

The two main factors in this equations are \( \alpha \) and \( \beta \), which represent the random initial status and the random rate of growth, respectively. These two main factors are the latent variables in the growth model. The growth term is multiplied by a coefficient, \( \lambda_k \), representing intervals of time and is typically a constant where \( \lambda_1 = 0, \lambda_2 = 1, \lambda_3 = 2 \), and so on. The unobserved error of measurement is represented by \( \varepsilon_k \).

The equations for the two factors of interest are:

\[
\eta_{\alpha i} = \mu_{\alpha} + \zeta_{\alpha i} \quad \text{for the initial status}
\]

\[
\eta_{\beta i} = \mu_{\beta} + \zeta_{\beta i} \quad \text{for the growth}
\]

where \( \mu_{\alpha} \) and \( \mu_{\beta} \) are the means of the initial status and growth rate, respectively, while each \( \zeta \) represents the residual terms associated with their particular endogenous variable. Combining the intercept and slope equations with the trajectory equations yields a combined model with both fixed and random components known as the reduced-form equation (Bollen & Curran, 2006):

\[
y_k = (\mu_{\alpha} + \lambda_k \mu_{\beta}) + (\zeta_{\alpha i} + \lambda_k \zeta_{\beta i} + \varepsilon_k)
\]

Translating the trajectory equation to a path diagram and adopting the conventional nomenclature (Jöreskog, 1970), Figure 1 presents the basic growth model.

This unconditional model can be extended to incorporate exogenous (independent) predictors for the growth and initial status factors. The matrix form of the intercept (Equation (2.1)) and slope (Equation (2.2)) equations is:

\[
\eta = \mu_{\eta} + \zeta
\]

where \( \eta \) is \( k \times 1 \) vector of latent factors (e.g. \( \eta_{\alpha} \) and \( \eta_{\beta} \)) and \( \mu_{\eta} \) is a \( k \times 1 \) vector of latent variable means. Including predictor variables to the model results in an expansion of Equation (4) into:
where $\Gamma$ is a $k \times p$ matrix of fixed regression parameters, $p =$ number of predictors in $x$. Similarly, extending the visualisation in Figure 1 to include two predictors ($x_1$ and $x_2$) is presented in Figure 2. Here we see that the latent factors ($\eta_\alpha$ and $\eta_\beta$) from Figure 1 become latent dependent (or endogenous) variables that are regressed on $x_1$ and $x_2$. These exogenous predictors are therefore explanatory variables for the initial status of the process ($\eta_\alpha$) and the rate of growth ($\eta_\beta$).

However, this conditional model is treating the covariates as time invariant (because they were only measured once) and as error-free indicators that are supposed to represent an underlying factor (Kline, 1998). This is a limitation that can be addressed by extending the conditional model to transform the predictor variables into more accurate latent factors, adapting an approach by Walker, Acock, Bowman, and Li (1996), and applied to a framework similar to simultaneous growth models (Curran, Harford, & Muthén, 1996). Instead of using a single measure of PS ability and using that indicator to predict both initial status and growth of math achievement, PS ability can be treated as a

\[ \eta = \mu + \Gamma x + \zeta \]
latent factor and modelled for growth. The latent initial status and latent growth for PS ability can be used as covariates for the respective latent variables for math achievement. This extension into a simultaneous growth model is described further in the methodology.

5. Methodology

5.1. Participants
Participants in this study are government school students who participate in the ALP testing every October and March. The ALP tests are designed to be administered to grades 3–10 students across the participating Department of Education and Early Childhood Development (DEECD) regions in Victoria. Testing commenced on October 2009 and was administered twice a year (March and October) every year thereafter. However, this study is focused only on grade levels 3–8, and only for participants from 2010 to 2011 test administrations. The participants for the ALP project within the scope of this study are 61 government schools in 6 DEECD regions in Victoria involving around 5,886 students of diverse linguistic backgrounds and a wide range of English-language proficiency. The sample characteristics follow the school population characteristics in terms of gender, language background and other demographic data that are collected by schools in Victoria. The sample
distribution is also reasonably close to state and national distributions for the general population within the age range of this study (for detailed demographic tables, see Vista, 2013).

5.2. Instruments and methods

The main instruments are the numeracy, literacy and PS tests from the ALP project (Assessment Research Centre, 2011). The development process for the ALP tests and the validation analysis (with NAPLAN as the validating measure) were described in detail elsewhere (Vista, 2013). As an overview, the tests are multiple-choice type, Adobe Flash™-based and administered online. The PS test items fall into three broad groups based on the type of reasoning process involved: spatial, symbolic or verbal reasoning. These items were matched to the Victorian Essential Learning Standards (VELS) under the general capabilities domain (Victorian Curriculum & Assessment Authority, 2012). The numeracy items have the following broad content areas: number, geometry, measurement, chance and data. The literacy items primarily measure RC skill (used in this study as the main RC measure) and were developed based on a previous large-scale literacy assessment study (see Assessment Research Centre (2004)). The ALP tests are in a single-scale calibrated using the Rasch model. As such, the “scores” reported are actually weighted likelihood estimates (WLE) of student ability in logits (centred at 0 with a standard deviation of 1). The numeracy and literacy tests were validated using the nationally administered NAPLAN (for details, see Vista, 2013).

Results show significant correlations between ALP and NAPLAN scores within same test administrations, implying concurrent validity. Results also show that when same cohorts were administered the ALP and NAPLAN tests two years apart, the scores correlated significantly, implying predictive validity. Regression analysis further supported this finding of predictive validity from the correlation analysis (see Vista, 2013). The measurement reliabilities of all subtests for Num, PS and RC test measures are reported in Table 1, with colours indicating the targeted ability levels for each subtest arranged in order of relative difficulty (easy to difficult from top to bottom). The test reliability measures used are Cronbach’s α and EAP/PV index, or the average expected a posteriori/persons variance, which

<table>
<thead>
<tr>
<th>Test</th>
<th>2010 Time 1</th>
<th>2010 Time 2</th>
<th>2011 Time 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EAP/PV</td>
<td>Cronbach’s α</td>
<td>EAP/PV</td>
</tr>
<tr>
<td>PS—Yellow</td>
<td>0.82</td>
<td>0.82</td>
<td>0.74</td>
</tr>
<tr>
<td>PS—Green</td>
<td>0.82</td>
<td>0.82</td>
<td>0.79</td>
</tr>
<tr>
<td>PS—Blue</td>
<td>0.85</td>
<td>0.85</td>
<td>0.74</td>
</tr>
<tr>
<td>Numeracy—Red</td>
<td>0.86</td>
<td>0.84</td>
<td>0.83</td>
</tr>
<tr>
<td>Numeracy—Orange</td>
<td>0.81</td>
<td>0.79</td>
<td>0.81</td>
</tr>
<tr>
<td>Numeracy—Yellow</td>
<td>0.83</td>
<td>0.82</td>
<td>0.82</td>
</tr>
<tr>
<td>Numeracy—Lime</td>
<td>0.80</td>
<td>0.79</td>
<td>0.82</td>
</tr>
<tr>
<td>Numeracy—Green</td>
<td>0.74</td>
<td>0.75</td>
<td>0.77</td>
</tr>
<tr>
<td>Numeracy—Aqua</td>
<td>0.68</td>
<td>0.74</td>
<td>0.77</td>
</tr>
<tr>
<td>Numeracy—Blue</td>
<td>0.79</td>
<td>0.71</td>
<td>0.72</td>
</tr>
<tr>
<td>RC—Red</td>
<td>0.81</td>
<td>0.85</td>
<td>0.77</td>
</tr>
<tr>
<td>RC—Orange</td>
<td>0.80</td>
<td>0.80</td>
<td>0.77</td>
</tr>
<tr>
<td>RC—Yellow</td>
<td>0.78</td>
<td>0.80</td>
<td>0.77</td>
</tr>
<tr>
<td>RC—Lime</td>
<td>0.79</td>
<td>0.78</td>
<td>0.77</td>
</tr>
<tr>
<td>RC—Green</td>
<td>0.77</td>
<td>0.77</td>
<td>0.77</td>
</tr>
<tr>
<td>RC—Aqua</td>
<td>0.77</td>
<td>0.77</td>
<td>0.77</td>
</tr>
<tr>
<td>RC—Blue</td>
<td>0.79</td>
<td>0.80</td>
<td>0.77</td>
</tr>
</tbody>
</table>
is the proportion of the estimated model explained variance to the total (Wu, Adams, Wilson, & Haldane, 2007). The component instruments for Num, PS and RC tests all have good reliability values.

The SEM approach adopted for this study is strictly confirmatory and concerned mainly in the confirmation of a predetermined structural model. The structural models being tested are simple unconditional latent growth models and conditional models with well-defined covariates, following frameworks that have been well established in the literature (e.g. Bollen and Curran (2006), McArdle and Epstein (1987), Meredith and Tisak (1990), Muthén (2004)).

The research questions were statistically tested by constructing models that were constrained to control specific model components. By comparing these constrained models with freely estimated models, null hypotheses on those specific model components can be tested for statistical significance. For example, in testing for linear growth, the baseline model will have no restrictions while the constrained model will have path coefficients that reflect a linear trend. If the fit of the constrained model is acceptable, this would suggest that there is no evidence for us to reject the model where linear growth is assumed.

5.2.1. Simultaneous growth model
Latent growth for PS ability was modelled using three test administrations for PS ability, matching with the three timepoints for the Num test administration. Each test administration is 1 semester apart (timepoint 1 = 1st semester 2010, timepoint 2 = 2nd semester 2010 and timepoint 3 = 1st semester 2011). Linear growth for PS ability was not the main focus of this analysis so the parameters of path loadings from the latent PS ability growth variable were not constrained as linear. Further, due to evidence that RC skill may be partially mediating the association between PS ability and growth in math (see Vista, 2013), this mediating effect is included in the extended model. Modelling predictors that are latent variables themselves is straightforward and involves using the latent variables of a growth model to predict the latent variables of another (see, for example, Kline (1998), Muthén (1991)). Thus, the latent initial status and growth variables of an LGM for PS ability became predictors for both latent variables in the growth model for math achievement. This model (designated as LGM-E1) is visually presented in Figure 3.

5.2.2. Dealing with missing data
Latent growth modelling was done using AMOS, and missing data were dealt with by computing maximum likelihood estimates through a procedure called full information maximum likelihood (FIML) (Anderson, 1957). Simulation studies done by Wothke (2000) show that, when applied to growth curve modelling, FIML is superior to pairwise and listwise deletions by providing parameter estimates with the least bias. FIML also tends to produce the least bias among comparative methods used in SEM (Arbuckle, 1996; Enders, 2001; Enders & Bandalos, 2001). More important, apart from producing unbiased estimates in missing completely at random (MCAR) and missing at random (MAR) types of data missingness (Rubin, 1976), FIML procedure is also robust to slight deviations from maximum likelihood estimation assumptions (e.g. multivariate normality) (Arbuckle, 1996). The type of missing values in this study is a scale-level missingness (e.g. students might take the Num and Lit tests but not the PS test). Further, since the subsequent administrations included new participating schools and there is no way to predict in the initial test administration which schools will decide to participate in the future, this type of missingness is closer to what is defined as MAR³ rather than MCAR or missing not at random (MNAR) (refer to Little and Rubin (1987) for details on the mechanisms of missingness).
6. Results

6.1. Main LGM results

The basic unconditional model for latent growth is presented in Figure 4. The two latent variables are defined as representing initial status and growth. These latent variables in turn load into the observed variables, which measure the levels of math achievement in three time periods. The path loadings from the initial status latent variable are fixed to 1, indicating that this latent variable is an intercept variable that sets the initial level of math achievement for the model. To test for linear growth, the default baseline model freely estimates the regression weight from the latent growth variable to the observed time 2 variable (denoted in the path diagram as $b_2$), while a second model constrains this regression weight to 1. The baseline model is saturated (i.e. zero degrees of freedom),
however by fixing $b2 = 1$, the constrained model can now be tested using model $\chi^2$ among other fit indices. If this model fits, meaning that regression weights of 0, 1 and 2 are acceptable, this implies that the latent growth variable is a linear parameter representing an amount of increase that is uniform over time (Duncan et al., 2006).

Comparative fit indices showed that the constrained model has good fit, $\chi^2(1) = 10.08$, RMSEA = 0.039, $P_{(RMSEA<0.05)} = 0.75$, CFI = 0.997, suggesting that linear growth is tenable. The parameter estimates for this linear model are presented in Figure 5. The mean initial status is 0.36, indicating that the average 2010 time 1 WLE estimates of student ability in math is 0.36 logits. The variance for the initial status mean is 0.60, SE = 0.03, $p < 0.01$, suggesting that there is statistically significant variation in the initial math achievement levels across all participants. The average growth is 0.26 logits per time period, again with significant variation of 0.03 across all participants, SE = 0.01, $p = 0.02$. Thus, the linear growth model based on these parameter estimates is $y = 0.36 + 0.26t$, for $t = 0, 1, and 2$. The implied means for each timepoint therefore are: Num 2010_1 = 0.36, Num 2010_2 = 0.62 and Num 2011_1 = 0.88. There is a negative and statistically significant covariance of −0.08 ($r = -0.61$) between initial status and growth, SE = 0.02, $p < 0.01$ suggesting that students with higher initial Num scores have a lower rate of growth, although this association has to be interpreted in the context of the small amount of variation in growth rates. This is unexpected, especially in comparison with results from studies that showed positive association between initial levels and subsequent
levels of performance (e.g. Lerkkanen et al. (2005)), but not contradictory since our results showed a negative association between initial level and the rate of growth in math achievement.

6.2. Simultaneous growth model
The simultaneous growth model did not achieve exact fit, $\chi^2(11) = 77.30, p < 0.01$, but is still tenable, RMSEA = 0.032, CFI = 0.993, and parameter estimates can still be interpreted with meaning. This allows us to examine the path loadings and the regression weights to answer the research questions.

Proceeding to evaluate the regression weights, latent PS ability growth does not significantly predict initial status in math, $B = 0.17, SE = 0.16, p = 0.29$. The direct effect of the mediating variable, RC skill, to growth in math is also nonsignificant, $B = 0.02, SE = 0.02, p = 0.32$. The covariance between initial status and growth for PS ability is nonsignificant, $\text{Cov}(PS_{\text{init}}, PS_{\text{growth}}) = -0.01, SE = 0.01, p = 0.26$.

For this extended model, the predictors account for 82% of the variability in initial level of math achievement and 53% of the variability in latent growth. Examining the factors that load on to these latent dependent variables, the main results with substantive relevance are the statistically significant predictors of the rate of growth in math achievement (Table 2). Students with higher initial PS
Table 2. Parameter estimates for the paths in Model LGM-E1

<table>
<thead>
<tr>
<th>Path</th>
<th>B</th>
<th>SE</th>
<th>CR</th>
<th>β</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>RC 2010 ← PS initial status</td>
<td>0.82</td>
<td>0.02</td>
<td>45.51</td>
<td>0.78</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Num initial status ← PS growth</td>
<td>0.17</td>
<td>0.16</td>
<td>1.05</td>
<td>0.05</td>
<td>0.29</td>
</tr>
<tr>
<td>Num growth ← RC 2010</td>
<td>0.02</td>
<td>0.02</td>
<td>0.99</td>
<td>0.09</td>
<td>0.32</td>
</tr>
<tr>
<td>Num growth ← PS initial status</td>
<td>-0.12</td>
<td>0.02</td>
<td>-4.95</td>
<td>-0.54</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Num initial status ← RC 2010</td>
<td>0.07</td>
<td>0.02</td>
<td>3.046</td>
<td>0.10</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Num initial status ← PS initial status</td>
<td>0.66</td>
<td>0.03</td>
<td>21.89</td>
<td>0.83</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Num growth ← PS growth</td>
<td>0.53</td>
<td>0.14</td>
<td>3.80</td>
<td>0.53</td>
<td>&lt;0.01</td>
</tr>
</tbody>
</table>

Figure 6. Non-standardised parameter estimates for LGM-E1. Statistically significant path loadings are highlighted in yellow.
scores tend to also have higher initial levels, $B = 0.66$, $SE = 0.03$, $p < 0.01$, but slower rate of growth, $B = -0.12$, $SE = 0.02$, $p < 0.01$, in Num scores. This association becomes slightly weaker if the partial mediating effect of RC skill is taken into account, reducing the total effect of latent initial status of PS to latent initial status of Num to $-0.10$ (indicating an indirect effect of 0.02). The rate of growth in PS ability also significantly predicts the rate of growth in math achievement, this time proportionally, $B = 0.53$, $SE = 0.14$, $p < 0.01$. This can be interpreted as such that, for every 1 logit increase from the mean growth in PS ability over the time period in this model, there is a corresponding increase of 0.53 logits in math ability. The mean rate of growth for PS ability is 0.18 logits, with statistically significant variability among the students in the data ($s^2 = 0.04$, $SE = 0.01$, $p < 0.01$). The latent means, covariances and path loadings are presented in Figure 6. The summarised means and variances are reported in Table 3.

7. Discussion

Preliminary analyses established that there is growth in Num scores across the three time periods that were observed. Data fitted a regression model where this growth is assumed to be linear. Fitting a latent growth model to the data resulted in good fit, with results indicating that, in an unconditional model, Num scores grow at an average of 0.26 logits per time period (i.e. around six months). The variation in the rate of growth is significant; some have faster rates of growth while others have slower rates of growth. A statistically significant variation is also present in the initial status of math achievement. There is statistically significant covariance between a student’s initial status and the rate of growth but, given that the variance in growth rates is small, the practical implications may not be substantial. After establishing that a latent growth model with linear growth parameters is tenable, the next step is including the predictors specified previously and to analyse whether or not they account for the latent growth.

This simultaneous growth model (LGM-E1) has good fit although relatively worse compared with the unconditional model. This could be due to added model complexity, given that this model has comparatively more estimated parameters and thus less parsimonious (Fan, Thompson, & Wang, 1999; Schumacker & Lomax, 2004). Because model fit remains good, the parameter estimates can still be meaningfully interpreted.

Initial PS ability predicts initial level of math achievement, supporting previous findings (Fuchs et al., 2010; Geary, 1994, 2011; Primi et al., 2010). This is expected, so we’re more interested in the dynamics of growth in math achievement and PS ability. Results showed that the average rate of growth in PS ability is 0.18 logits per time period. The magnitude of this rate of growth is almost double compared with the mean rate of growth for math achievement, which could be interpreted as an indicator that PS ability is more amenable to change than math achievement. This rate of growth significantly predicts the rate of growth for math achievement such that, on average, those who are growing in PS ability at a rate of 1.0 logit faster tend to grow in math achievement at a rate of 0.53 logits faster as well. This is somewhat compensated by the inverse association of latent initial status in PS and latent growth in math such that those who are 1.0 logit higher in initial PS ability
have a lower rate of growth in math by around −0.10 logits per time period (direct effect of −0.12 plus indirect effect of 0.02). But because the loading on math growth from PS growth is larger than the loading from PS initial status, those with higher rates of growth in PS ability will still end up with higher rates of growth in math regardless of their initial PS scores.

The results from the extended model also allow for the comparison of the growth rates for both math achievement and PS ability. According to the model, students have an average rate of growth in PS ability that is higher than their rates of growth in math achievement (\( M_{\text{Num}} = 0.10, \ SE = 0.05, \ p = 0.03 \) vs. \( M_{\text{PS}} = 0.18, \ SE = 0.04, \ p < 0.01 \)). Interestingly, there is no covariation between latent initial status and growth of PS ability, but the variance in PS ability latent growth is significant (see Table 3). This suggests that, for whatever reason, one’s growth rate in PS ability does not appear to be associated with one’s level of PS ability in the beginning. Whatever is causing some rates of growth in PS to be higher than others is outside the model. This finding could have important implications for pedagogy and the developmental framework approach to learning, to be discussed in more detail later on. The association between PS ability growth rate and trajectory in math achievement is presented in Figure 7. Here we can see that, as the growth rate in PS ability increases, the growth rate in math achievement also increases (i.e. the slope of the trajectory in math achievement over time becomes steeper).

A simplified plot that takes into account the effect of initial PS ability status is presented in Figure 8. In this plot, trajectories for three levels of initial status are shown, each with a specific rate of PS ability growth in relation to the mean values for both latent variables. Here we see that those who have an initial PS ability of 1 SD below but with a rate of growth 1 SD above the mean (low-start, high-growth or LSHG group) is catching up in terms of math achievement with those who started at
The results also show an indirect effect of PS ability on math initial status of around 0.02 that can be attributed to RC skill. There is no evidence of any similar mediating effect on the latent growth variables. It is possible for a student’s initial RC skill to be mediating between PS ability and math achievement early on, although the effect is weak. This is not surprising, as it is certainly possible that a portion of the latent difficulty in the math items may be attributable, but not exclusively, to reading loads (Guglielmi, 2012; Jordan et al., 2003; Landerl, Bevan, & Butterworth, 2004). In particular, Guglielmi’s results based on large-scale data suggest that RC skill may be contributing to math achievement independent of general reasoning ability (Guglielmi, 2012). However, the nonsignificant path loading from RC skill to latent growth in math suggests that over time, as the student becomes more immersed in the classroom, initial levels of RC skill may have increasingly less effect on subsequent progress in math achievement compared with the influence from PS ability.

From a pedagogical perspective, this does not lessen the importance of improving RC and related language skills among students. On one hand, there is evidence in the main LGM model that RC skill covaries with PS ability. This highlights the need to focus more on language skills especially for NESB students because the teaching of reasoning skills in Australia is essentially carried out in English. On the other hand, the results show that while RC skill is associated with initial status in math, growth in math achievement becomes independent of initial RC skill from that moment forward. This implies that the disadvantage in initial math achievement due to lower levels of RC skill can be compensated by other factors. In addition, improving the PS ability of students can lift the math achievement of those who have relatively lower initial levels and allow them to catch up.

This result confirms findings from the literature that math learning involves considerable general reasoning ability (e.g. Casey et al. (1992), Hart et al. (2009), Primi et al. (2010)). However, the strength of the association between PS ability and math achievement is not uniform longitudinally such that the amount of variance in math growth being explained by PS ability (and RC skill) decreases in
magnitude over time for any individual regardless of their grade level at the start of this study. This means that time as a variable tends to have a negative effect on growth in math achievement. Because student year is related to levels of PS ability (i.e. those in the higher grade levels have higher PS scores), this effect is captured in the latent growth models such that PS ability has a negative loading on the latent growth factor for math achievement. The causes for this effect of time on the trajectory in math achievement are complex and may await future analyses with more comprehensive data on both within and between-individual factors. For example, future data on classroom changes over the course of a student’s school life may provide insight on why growth appears to slow down in the secondary levels relative to early primary levels. Curricular shifts, changes in student motivation, even cognitive developmental maturation may all play very important roles in further explaining the shifts in math growth.

In contrast to this time effect on math growth, growth in PS ability does not appear to be related to initial levels. It can therefore be argued that growth in PS ability is less affected by time-related factors compared with growth in math achievement. Another interesting result is that the average rate of growth for PS ability is almost double that of math achievement. This could have important implications for pedagogy because general reasoning skills have long been found to have strong associations with math learning (McGrew & Hessler, 1995; Taub et al., 2008) but have only been recently included as a formal component of the curriculum in Australia. Once firmly established in the curriculum, targeted teaching of reasoning skills may be able to lift or enhance math learning if we follow the logical consequences of experimental evidence based on such interventions (Nunes et al., 2007).

7.1. Implications and future directions
The growth models allow the trajectory of math achievement to be extrapolated beyond the time span in this study within a certain reasonable extent. While both regression and latent growth models suggest an eventual levelling off in the trajectories of those who have high initial levels of PS ability, the models also imply that manipulating the rates of growth in PS ability can have a direct positive impact on math achievement growth.

These findings have important implications for teaching under a developmental framework where assessment data are used primarily to improve learning regardless of the initial status in a particular domain of learning (Griffin, 2007). When applied to the teaching of problem solving and general reasoning skills, the initial status in PS ability would indicate the initial level of a student’s readiness to learn (Griffin, 2007) so that teachers can scaffold content and approaches for optimal learning. This study shows that facilitating a higher growth rate in PS ability can translate into increased math achievement over time, regardless of initial status in either math or PS ability. In the extended latent growth model, predictors of math growth account for as much as 53% of the variance in latent growth in math achievement, with PS ability growth rate as the strongest predictor.

The factors that account for the growth rate in PS ability are beyond the scope of this study and thus were not examined here. These could come for example from home environment changes or developmental maturity, as well as from a diversity of other factors. However, it is natural to consider that growth in PS ability may be facilitated through school intervention. A study by Nunes et al. (2007) found that training children in logical reasoning can improve their math learning more than children who were not given logical reasoning instruction. Their results are confirmed in this study’s findings where growth in PS ability predicts growth in math achievement. The implications could be far-reaching because if we relate the findings in this study with the findings from Nunes et al. (2007), the beneficial effects of growth in PS ability through specific intervention could last for a substantial length of time and affect math learning across a wide area of the curriculum even if the general reasoning instruction does not target any specific math domains.
7.2. Limitations and recommendations
The main limitations of this study regarding scope (government schools in Victoria), range of academic year levels included in the study, and management of missing data are discussed in an earlier and related publication (Vista, 2013).

Specific limitations include the relatively low number of timepoints in this study. Three timepoints enabled testing for linear growth but more timepoints over a longer period of time would provide more data to describe the latent growth curves better. Extending the study has logistical implications as well as methodological challenges, especially concerning the heightened chances of missing data. Perhaps a balance between sample size and length of study could be considered for future study designs. Longer duration and more timepoints also need to be weighed logistically with the option to include additional independent cohorts for more robust cross-validation.

It is also recommended that the trajectory models in this study be validated by future research. Independent samples with either similar or different scope in terms of sample characteristics will allow a theoretical validation of the growth trajectories that were fitted to the data in this study. For example, country-level representative samples, samples that include a wider age-range or cohorts based on longer time-spans are recommended. The inclusion of other demographic variables in the models, such as SES or parental education, may also be useful. These validation studies can help strengthen the findings or, if future results are contradictory, provide a basis for re-examination of the conclusions put forward here.

Finally, it is hoped that the issues and challenges tackled in this study provide some insight to future studies of similar nature. It is recommended that the results and implications find their way to policymakers so that they can be translated into operational terms. The research possibilities on the dynamics of math growth and factors that affect it remain exciting. It is up to future researchers to validate the findings as well as extend this study to look at growth dynamics in other areas of student learning.

Funding
The author declares no direct funding for this research.

Author details
Alvin Vista1
E-mail: alvin.vista@acer.edu.au
1 Melbourne Graduate School of Education, University of Melbourne, Victoria, Australia; Australian Council for Educational Research, Camberwell, Australia.

Citation information
Cite this article as: The role of PS ability and RC skill in predicting growth trajectories of mathematics achievement, Alvin Vista, Cogent Education (2016), 3: 1222720.

Notes
1. NAPLAN scores are also on a uniform scale that spans from year 3 to year 9, and represent the same ability level over time (Australian Curriculum, Assessment and Reporting Authority, 2011).
2. That is, the instruments composing each test before they have been horizontally and vertically calibrated into a single scale under the Rasch model.
3. This is only an assumption because we can only test whether or not the missing values are MAR and not whether they are either MAR or MNAR (Newman, 2003; Schafer & Graham, 2002).
4. Because growth rate is not associated with initial status in PS ability, this graph is fixed at the mean level for initial status. In other words, this graph shows the association for those with the mean initial status of PS ability (\(M_0 = -0.36\)).
5. Under the thinking processes strand of the AusVELS (Victorian Curriculum & Assessment Authority, 2012) and as critical and creative thinking component of the seven general capabilities in the Australian Curriculum (Australian Curriculum, Assessment & Reporting Authority, 2012). Both curriculum frameworks are still in development as of 2012.

References
Psychology, 45, 850–867. 
http://dx.doi.org/10.1037/0014-9059
http://dx.doi.org/10.1093/biomet/57.2.239
http://dx.doi.org/10.1207/s15327043hup1803_3
http://dx.doi.org/10.1007/BF03173314
http://dx.doi.org/10.1016/j.cognition.2003.11.004
http://dx.doi.org/10.1016/j.lindif.2010.05.001
http://dx.doi.org/10.1007/BF03173503
http://dx.doi.org/10.2307/1130295
http://dx.doi.org/10.1146/annurev.psych.60.110707.163612
http://dx.doi.org/10.2307/1130295
http://dx.doi.org/10.1177/073428299501300102
http://dx.doi.org/10.1007/BF02294746
http://dx.doi.org/10.1177/1094428103254673
http://dx.doi.org/10.1037/0022-0663.97.3.299
http://dx.doi.org/10.1016/j.lindif.2010.05.001
http://dx.doi.org/10.2307/2527726
http://dx.doi.org/10.1080/01443418700707107
http://dx.doi.org/10.1093/biomet/63.3.581
http://dx.doi.org/10.1037/1082-989X.7.2.147
http://dx.doi.org/10.1111/jd rp.2006.21.issue-4
http://dx.doi.org/10.1093/acprof:os o/9780195125968.001.0001
http://dx.doi.org/10.2307/749636
http://dx.doi.org/10.1037/a0013830.23.2.187


