Appendix 10 Individual learning trajectories.

The data analyses developed in Chapters 6 and 7 do not use longitudinal data for the observation of learning growth of individual students but snapshots at one point in time for a large set of individuals across 8 Year levels. As indicated in Chapters 5 and 6, the general trends in these cross sectional means of learning growth with Year level mostly approximate the general trend shown with the means of longitudinal data. Longitudinal data for individual students is another matter. Each trajectory is unique. As a consequence, teacher generated assessment data through observations will require different processes of recording and analysis than those conventionally applied in classrooms. This appendix considers some of the issues involved in individual student trajectories.

Whether teacher judgement data become an additional data source or not, schools will need to develop (or adopt) processes to better record and analyse student learning growth. There is pressure to do this generally and in particular from US policy analysts (Duncan, 2009; Rudner & Boston, 2003; Smith, 2008; State Educational Technology Directors Association, 2008). Unsurprisingly software and system suppliers (Ligon, 2009) also advocate the development of data warehouses and reporting process for individual student growth trajectories. The pressure is for better access to data and data driven decisions but "not advocating for additional high-stakes tests, instead … that using technology to assess students in a less formalized, yet more personalized, manner can glean benefits for teachers and students alike" (State Educational Technology Directors Association, 2008, p. 1). One source for this data will be computer adaptive testing. Another potential source is general classroom assessments, including teacher judgement assessments.

There is debate about how feasible it is to use summative, formative and interim assessments, to serve multiple purposes. Interim assessments are "assessments that fill the gap between classroom formative assessments and state summative assessments … an integral part of any comprehensive assessment system and should be evaluated as such" (Perie, Marion, & Gong, 2007, p. 1). However it is consistent with the concept of data warehousing, seen as underpinning the access to data and data mining, that as wide a range of assessment data as is possible is archived for a student. Were the data able to be stored using common scales across all assessments, it seems logical that these could generate a set of time related data points for each student.

As part of the thought experiment the nature of the data potentially available is considered. A data point could be stored with a minimum of five elements. These would be the student identifier, the strand of the curriculum, the source of the assessment, the assessment value and an automatic time stamp. Where data were not automatically recorded on the appropriate common scale, the source of the assessment might provide a conversion protocol to that scale. The student identifier provides the link to additional information about the student. Student and strand together provides a link to the class identifier. Entries could be automated (particularly if from other systems - adaptive testing, state tests etc.) or designed to minimise data entry requirements. One minimising control for teacher entries might be a policy of adding a new point for a student only when new development by the minimum scale unit has been observed. New wireless technologies have already been used to advantage to simplify teacher record keeping (Wireless Technologies, http://www.wirelessgeneration.com/).

Assuming 10 or so data points per year per student (per strand of the curriculum), how might this data be analysed within a year for the current teacher, and over a longer time scale (up to the whole school life of a student) for the benefit of a student and the other teachers with responsibility for that student? The question of when the student record would need to be destroyed is not discussed here but it is noted that the recording and/or the extended preservation of schools grades, as they would be seen, raises a significant 'privacy' issue.

If the trajectories of learning for individual students were to be made visible to teachers and students (through graphical presentation), what might the data and images look like? A visual representation of the data points for individual students showing the trajectory to the present position would provide a teacher with an understanding of the current status (a position with meaning; x=likely to be able to do this, unlikely to be able to do that) and the recent and previous rates of learning (the gradient with time from earlier points on the scale). This image in itself might be sufficient for a teacher. In principle, it might be feasible to enrich the teacher's understanding with estimates of likely progress points into the future. This appendix explores in general terms what individual trajectories of learning look like, based on available public data sources and the extent to which helpful forward projections of trajectories might be feasible.

Overview of the sources of longitudinal data for individuals

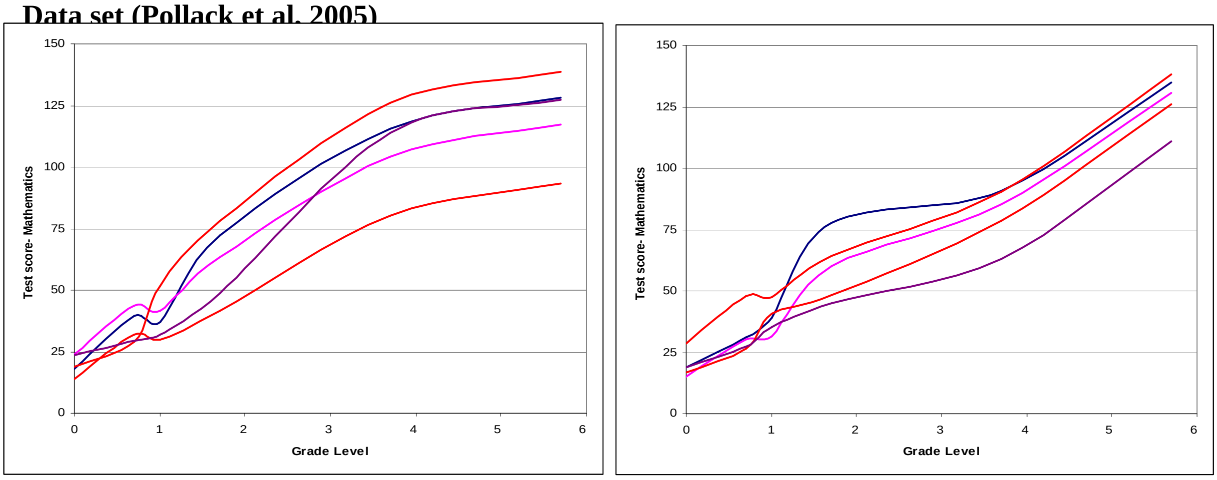
Data from the Early Childhood Longitudinal Study (ECLS) (Pollack, Atkins-Burnett, Najarian, &Rock, 2005) illustrate what individual trajectories look like for Year levels K to 5. Some analytical issues that might need to be incorporated into computer processes to support teachers with understanding and interpreting individual student trajectories are then considered. Examples of large longitudinal data sets are not readily found and many current initiatives that generate and manage individual student data seem to have moved into commercial products. As such, they are often protected from public access to data and processes. Contemporary approaches and analysis techniques for time recorded learning data built into commercial products were not found in the literature, although reviews of the products themselves are available (Ligon, 2009; What Works Clearinghouse, n.d.).

In the 1970s the issues being addressed in student trajectories in a then developing computer assisted learning project, were regularly published in the psychological literature. The early work of the Stanford University and Computer Curriculum Corporation (CCC) in the 1960s and 1970s, before it too became 'commercial in confidence', provides some understanding of the then thinking of how longitudinal records might assist in the management of learning.

Public examples of individual pathways

Figures A10.1 and A10.2 are panels of idealized learning trajectories taken from the US ECLS database. The trajectories are smoothed and are based on 6 points only (two in K, two in Grade 1, and single points in Grade 3 and Grade 5). Their purpose is to illustrate the wide variation in the pattern of pathways taken by students who start close together and may even and arrive at approximately the same point or vastly different points after 6 years. Figure A10.1 shows two sets of trajectories for Mathematics. The left panel illustrates students who start fast and rise to a score of 100 or above by Grade 3. The right panel shows a second group who have started slowly and then accelerated from Grade 3 to Grade 5. One case in the right panel shows a quick rise to a score of above 75 by Grade 1, no growth from there to Grade 3, and then high growth from Grade 3 to 5.

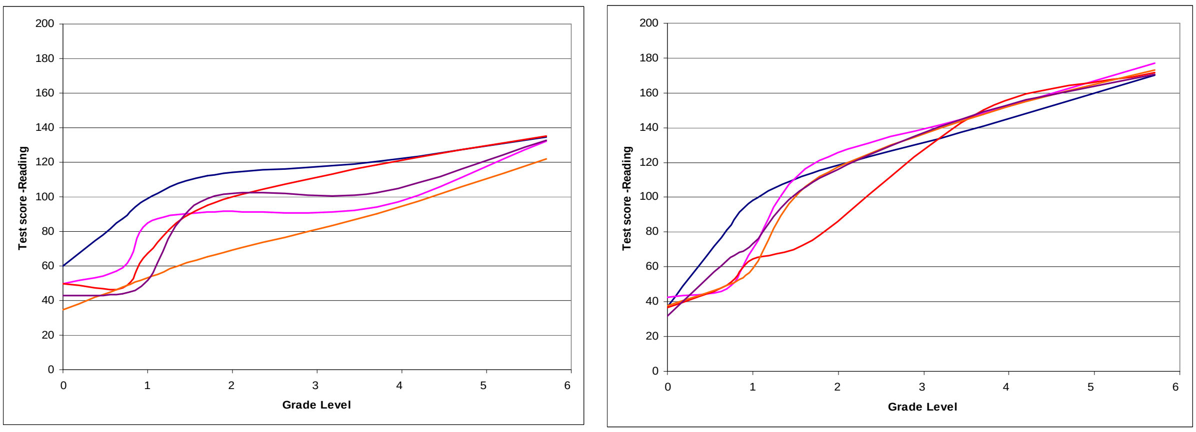
Figure A10.1 A sample of individual trajectories in Mathematics from the ECLS Public Data set (Pollack et al, 2005)



The panels in Figure A10.2 show another view of the destinations after 6 years, in this case for Reading. The left panel students start in a range from 38 to 60 and grow to 120. In the right panel the students start around the 40 score region, and via different pathways (similar to Figure A10.1) grow to around 170.

The purpose of the examples is to indicate the wide variability in trajectories and the complexity this raises for teachers in anticipating what might happen next and what intervention might be beneficial for each student. It is assumed a teacher would find a range of tools useful in dealing with this data, one part of the Fullan at al. Knowledge Base. Tools would include processes for estimating learning status (learning progressions), processes for visualising histories and plotting trajectories and indicators of what to do next at particular scale points. Examples that show the slow-fast-slow, steady or highly variable trajectories that students follow could be provided to help teachers identify outlier cases.

Figure A10.2 A sample of individual trajectories in Reading from the ECLS Public Data set (Pollack et al, 2005)



The data illustrated are very 'smoothed'. As data become available at shorter and more frequent intervals the complexity for teachers in making sense of the data will increase. This will be partly as a result of the increased variations around the 'true' trajectory due to measurement error. The reasons for the variations in trajectories not due to measurement error are beyond the scope of this thesis. Some recent research projects (Parrila, Aunola, Leskinen, Nurmi & Kirby, 2005; Aunola, Leskinen, Lerkkanen, & Nurmi, 2004; Hoeksma &

Kelderman, 2006) have applied Structural Equation Modelling, Hierarchical Modelling and Latent Class analysis to identify some of the contributing issues.

Making the trajectories visible

The visualisation of data is seen as a helpful process in its own right. At an earlier time in the apparently simple area of monitoring height, Burgess (1937) observed that many of the apparent irregularities of growth were really due to carelessness in measurement.

> The rate at which a child is growing is beginning to be regarded as one of the important indicators of his general physical condition. … Many school and private physicians who watch weight very carefully are content to measure height to the nearest inch often without regard to posture, or to measure with shoes sometimes on and sometimes off so that, according to their record; children apparently shoot up or shrink down in the most startling fashion. In many records, especially for younger children, heights and weights are transposed on the record card and much interpretation is needed to get the height picture approximately accurate. (Burgess, 1937, p. 305)

Burgess however makes the point, which the author believes will also apply when teachers have richer regular learning data (much from their own assessments and observation) and can see the trajectory of individual growth, that this will be its own incentive to develop better data procurement processes and to recognise where double checking will be required. What Burgess anticipated was that regular recording and graphing of height information would bring its own insights into error and irregularity, and alert observers to any anomalies.

> When a height chart is kept for the individual child, and his height line entered month by month or term by term an error in measurement stands out as dramatically as does a wrong thermometer reading on a fever chart. Where measurements are verified but the child suddenly begins to grow at an abnormal rate, either much faster or much more slowly than is usual, the graphic record gives parents, teachers and physicians prompt warning that he needs to be kept under close observation, and possibly given special medical care. The physician does not of course, make a diagnosis based on height alone, but a careful growth record is often a valuable diagnostic aid. (Burgess, 1937, p. 306)

The importance of height development per se may be less in the current context of generally better nutrition for children, but Burgess's insights about how serious readers of data react to anomaly are apt. Can a possible future where teachers react in the described manner to learning data be anticipated? Could individual development records on common scales help individual learning? A complicating issue for teachers is the uniqueness of each individual, trajectory, even though as illustrated in Chapter 5 the mean trajectories of groups of students can be modelled.

Each individual development trajectory is unique.

Since each individual trajectory is unique, to what extent can the patterns for groups, the average trajectory of the group with time, help in the understanding of individual trajectories? Keats (1983) cited by Willett & Sayer (1994), deemed models as having the property of *dynamic consistency* when the curve of the averages is identical to the average of the curves. Where dynamic consistency does not apply, the character of the individual curves is unrelated to the group trajectory, making it difficult to infer the shape of individual growth from a group growth curve. The variability in individual pathways illustrated in Figures A10.1 and A10.2 suggest that predicting an individual path at any time will be open to considerable uncertainty.

While increasingly sophisticated models for analysing change with time are available (Singer & Willett, 2003; Collins, 2006; Cudeck & Harring, 2007), models developed in this thesis for test data and teacher-assessed data are rather simple, particularly since the cross-sectional views are snapshots of time. It is impossible, for the author at least, not to wonder what the trajectories of students had been up to the point of the snapshot and where they might be at future points in time. This wondering raises the broad issue; can a 'black box' be developed for teachers to help them understand the implications of the pathways for each student and how to maximise their trajectories?

When processes for describing (modelling) individual student learning development over time within a class and across class years (grades) are considered, as might be required in such a 'black box' tool to help teachers with their decisions for appropriate types and timings of interventions, the feasibility of modelling individual trajectories needs to be addressed. How might this be done? What patterns might be expected? What might be the 'control boundaries' (in a quality control process) of development in say mathematics? When does a case change from being within expected ranges to being well outside? To what extent can group data assist with estimating 'safe' trajectories for individuals?

According to Molenaar (2004) "modern psychology is saturated with probability models and statistical techniques" (p. 202). However he believes that psychologists "attention is almost exclusively restricted to variation between individuals (interindividual variation [IEV]), to the neglect of time-dependent variation within a single participant's time series (intraindividual variation [IAV])" (Molenaar, 2004, p. 202).

He argues that most psychological processes should be considered as non-ergodic. The property of being non-ergodic implies a system that is influenced by history and is thus less predictable for lack of repetition of previous states. In contrast an ergodic system will return to states that are closely similar to previous ones.[37] Molenaar argues that the learning trajectory for an individual is non-ergodic. Furthermore, consistent with Keats's dynamic consistency, knowing the pattern with time of a population (IEV) does not necessarily assist in estimating the trajectory of an individual. In non-ergodic processes, an analysis of the structure of inter-individual variation will yield results that differ from results obtained in an analogous analysis of intra-individual variation.

> Hence for ... all developmental processes, learning processes, adaptive processes, and many more, explicit analyses of IAV for their own sakes are required to obtain valid results concerning individual development, learning performance, and so forth. (Molenaar, 2004, p. 202)

The essence on Molenaar's argument is that different approaches are required and different results are obtained when one follows individuals over time, as against aggregates of individuals. This point is made to support the complexity of the problem that faces teachers were more data provided to them, or developed by them, to follow the learning trajectories of individual students. Based on Molenaar's analysis any computer support system for the management of learning that depends upon simple extrapolations of individual trajectories from population patterns would be inaccurate. Recent publications (Molenaar & Campbell,

---

[37] Ergodic theory goes back to Boltzmann's ergodic hypothesis concerning the equality of the time mean and the space mean of molecules in a gas, i.e., the long term time average along a single trajectory should equal the average over all trajectories. The hypothesis was shown to be incorrect but the identification of a class of processes that have the property of tending to return to a previous state does provide a reference for considering 'converse' systems, particularly developmental systems.

2009; Molenaar, Sinclair, Rovine, Ram, & Corneal, 2009) indicate that Molenaar and colleagues believe there is still little literature and analytical support for non-ergodic cases:

> We are at the brink of a major reorientation in psychological methodology, in which the focus is on the variation characterizing time-dependent psychological processes occurring in the individual human subject. It will require substantial efforts from the community of psychological scientists to effectuate this reorientation. At present, there is very little literature on multivariate time-series designs and analysis techniques tailored to dealing with non-ergodic psychological processes. (Molenaar & Campbell, 2009, p. 116)

Molenaar et al. (2009) develop the argument for, and provide examples of, non-stationary time series modelling to address the problem of analysing individual level data.

> The EKFIS [extended Kalman Filter with iteration and smoothing] is a new and promising tool to analyse nonstationary time series in accordance with the classical ergodic theorems and with the basic tenets of DST [developmental systems theory]. Several aspects of the EKFIS are still under ongoing investigation, including alternative ways to determine the standard errors for the estimated time-varying parameters and technical aspects associated with the EM [expectation–maximization] loop in which the EKFIS is embedded as expectation step. Yet the results obtained thus far with the EKFIS indicate that it constitutes a viable and principled approach to the analysis of non-ergodic (nonstationary) developmental processes and thus allows for articulation of the basic tenets of DST—that individuals are complex dynamic systems, the characteristics of which are, themselves, changing and developing over time. (Molenaar et al., 2009, p. 369)

It is assumed that for data held on students from computer–adaptive testing, methods of analysis will be needed beyond the simple graphing of trajectories and summarising of norms especially where a reliable forward projection for an individual is expected. It is reassuring, for the author based on Molenaar's contemporary 2009 view, to appreciate that this problem is understood but not yet solved. The meagre literature search results appear to reflect that researchers are only at an early development stage for estimating trajectories for individuals. The provision of useful analytical tools and models of individual leaning growth, as required for teachers in the Fullan et al. Knowledge Base, will depend on further research.

Early attempts at individual based models of growth

There appear to be few sources for understanding individual trajectories as observed with small time increments. Of these few, a set of analyses come from the first major computer assisted learning projects, as a result of recording each student response. Starting in the 1960s a mathematical and practical approach to modelling of education development was explored and applied by Suppes and colleagues, based on work at Stanford University and the

Computer Curriculum Corporation (CCC)[38]. These models were developed to understand and predict trajectories for students involved in this early computer assisted instruction. In many cases the time dimension was defined by ‘number of trials’, a more precise measure than age or time on task. The materials were regarded as a behaviourist approach to learning development (Mazyck, 2002) but for the purposes of considering mathematical models of development, they provide example sets of individual trajectories and explorations of how they might be recorded and used for estimates of learning status at subsequent future time periods.

The two elements for generating a trajectory for learning for an individual are reliable measurement of learning and an adequate time metric. Where an IRT approach to measurement is applied, the measurement points on the time axis need to be spaced appropriately so that the change in learning over a unit of time is comparable to or less than the error of measurement for the learning status estimate. The frequencies of measurements, or the time intervals between measurements, affect the smoothness of the trajectory. Smoother, usually increasing trajectories of learning are generated when fewer measurements per unit time are applied (as shown in the smoothed example for ECLS in Figures A10.1 and A10.2 above). However, as each point is estimated with error, the more scale-readings the more likely the curve of best fit through the data points will reflect the ‘true’ trajectory. Most longitudinal studies cannot afford to measure at time increments below 6 months (for logistical reasons and because of the impact of testing in many cases on the students) and thus the number of points per individual is small, each measured with error. Data that show learning for individuals at small intervals between measurement points need less intrusive methods. Data derived from routine automated record keeping is one option for ‘embedded measurement’.

This was achieved in the initial development of computer assisted learning assessments by recorded points being ‘embedded’ as a result of the mastery learning process. Measurement in this case is different to the IRT model. Problems of a specific type were repeatedly presented to the student, changing the values in the problem, until a specified criterion of consistent correct responses was met. An assessment data point was created when the criterion was achieved. While pedagogical processes have moved on from this stage, and the approach is now offered as a supplementary process only (SuccessMaker .http://www.pearsonschool.com/, June 2009), the early years of the process generated a unique opportunity to observe the shape of learning with time.

Suppes, Fletcher & Zanotti (1976) developed a simple set of 5 axioms to explore what they termed “student trajectories rather than student progress, in order to give the sense of a definite path as a function of time that we are predicting for the individual student.” (p. 118).

The 5 axioms considered rates of processing, time effects, the effect of introducing new material, position (status) in a course (represented as a grade and progress in a grade in a decimal metric; e.g. 3.2 indicates grade 3, plus 0.2 of the way through grade 3 level) and the rate of progress in a course in relationship to the rate of introducing information in the course.

---

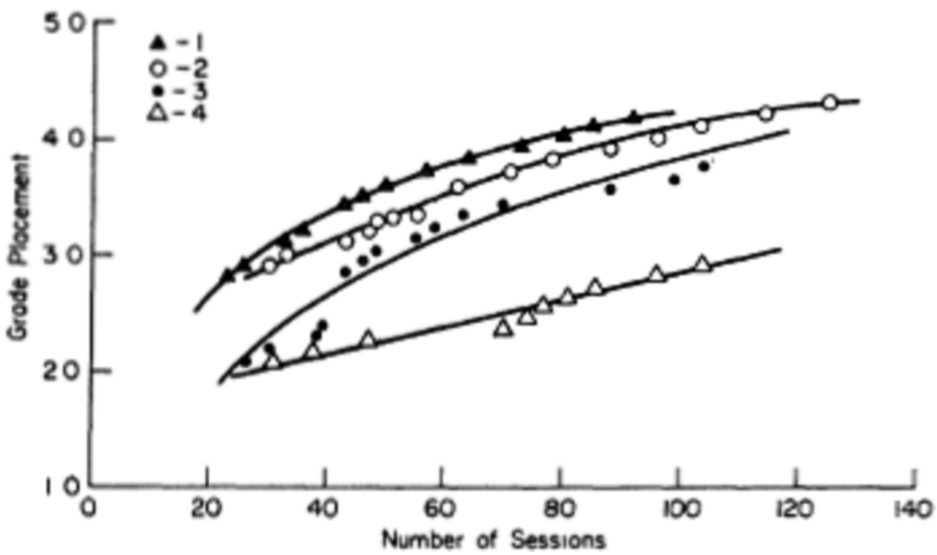
[38] Suppes founded the Computer Curriculum Corporation (CCC) in 1967. In 1990, it was acquired by Simon & Schuster, and then in 1999 by Pearson. One of CCC's major products was SuccessMaker. It is currently (2009) marketed by Pearson and includes 3,300 hours of supplemental instruction in English, language arts, math, science, and social studies in individualised, self-paced lessons, with the starting level individually determined and with diagnostic advice provided for recurring misunderstandings. ‘Forecasts tell you which students will meet instructional goals and when’. (Pearson website http://www.pearsonschool.com/, June 2009).

These considerations enabled Suppes et al.to develop some approaches to the general analysis of student trajectories in a maths curriculum covering roughly 7 years of elementary schooling. The curriculum was broken down into 14 parts, each corresponding to about half a year and included 14 strands (number concepts and decimals as two examples of strands) that were covered in some or all of the 14 time parts. Number concepts started in grade 1 and continued to grade 7.9, the only strand that occurred in all periods. Decimals, as an example, started in grade 4.0 and continued to grade placement 7.9.

A student's progress through the graded strands structure was a function of his/her own performance and was independent of the performance of other students. Progress on a given strand was also independent of performance on other strands. Movement through a strand used the pattern of correct and incorrect responses to insure a rate of movement that reflected performance. This structure has parallels with the general concepts of levels, strands and learning areas addressed in this thesis and described in Chapters 3 and 4. In particular, the individual progression of students meant that a student could be in a Grade 3 class but dealing with say grade 2.4 mathematics material, or for another student in the same class, material at grade 5.2.

Figure A10.3 below taken from Suppes et al.(1976), illustrates four typical cases from individual trajectories of almost 300 hearing-impaired students who participated in the program over a number of years. The grade placement value (GP) is the average of all strands for the student. The session number plotted was the one where the student had moved up (or down) 0.1 of a GP, achieved when students had worked through about 400 maths exercises, the actual number dependent upon error rates. This criterion for a plotted point ensures a relatively smooth curve as it reduces the error of measurement effect for points on the vertical scale that applies with IRT measurements.

Figure A10.3 Individual student trajectories- graphic from Suppes et al.(1976)



Three of the cases have lines of best fit that are clearly curves (1 to 3) while 4 is almost linear. Students 1 and 2 are very close after 20 sessions, but then follow different trajectories. Likewise for students 3 and 4 closeness after 20 sessions leads to quite different trajectories. Suppes et al.(1976) proposed that each trajectory could be described by the equation

$$y(t) = bt^k + c \qquad (1)$$

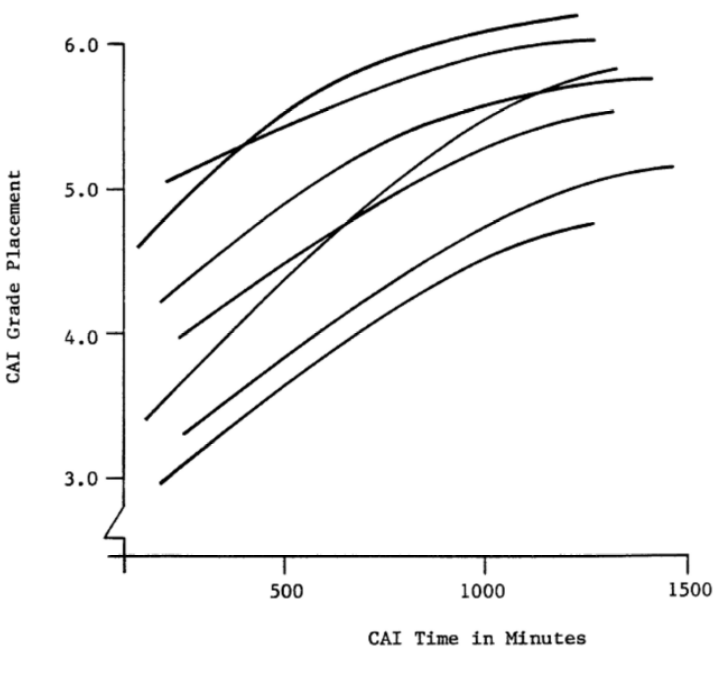
where y is the position (Grade Placement) at time t; b, c and k are parameters specific to the individual. While they describe the process as stochastic (as against deterministic) they do not include measurement error terms in the formulation (which is understandable given the basis for estimating GP discussed above).

The approach adopted was ground breaking in the 1960s and 70s and only a first exploration of the ways of modelling longitudinal learning data. Suppes et al.(1976) argued the need for a global theory of a student's progress through a curriculum (Suppes et al, 1976, p. 126).

Malone, Suppes, Macken, Zanotti & Kanerva (1979) developed 10 mathematical models based on 'power' functions of the general form of equation 1, for predicting a student's final grade placement. Data were obtained from 2000 elementary students at 2 weekly intervals for a full school year. Two simple models based on the most recent point and parameters estimated for the whole group, were the best for predicting end of year GP values. A power function model estimated individually for each student was best for describing all observed values for that student (using mean standard error over all students of data points compared to the fitted curves as the fit measure).

Later work in 1980 from the same general team (Macken, Suppes & Zanotti, 1980) argues against a global theory applied to the group without exploration of the patterns of learning over time (trajectories) for individuals. They presented Figure A10.4, making the point that while all the cases were one year below their chronological grade level each trajectory is distinct and a mean trajectory would not represent any one of the lines. This concern is similar to that raised more recently by Molenaar (Molenaar & Campbell, 2009).

Figure A10.4: Examples of individual student trajectories from Macken et al.(1980)



The individual trajectories can be described by the general equation (1) above only when the parameters b and c are estimated for the individual and not when they are estimated for the group. By implication, the estimates of the parameters for the individual are critical where judgements need to be made about whether an individual is performing outside the pattern that best describes their previous development.

A key insight is that the relationship between time and gain is not linear, even for individuals (or for grouped data as illustrated earlier). Macken et al. were concerned that evaluations of computer aided instruction (CAI) would misunderstand this point. Evaluations that assumed a linear relationship would mask some of the effects of individualised instruction. This thesis argues that, at a broader level, assessments of any students over time that are not sensitive to the trajectory of the individual will misinterpret when and what assistance might be applied to individual students, even where more useful tools for teacher assessment are applied. As summarised by Macken et al, ‘individuals proceed through the curriculum with distinct velocities and accelerations. The amount of gain per unit of time is different for different individuals and for the same individual across time.’ (p. 82-83).

Conclusions about individual longitudinal records

As teachers are encouraged to connect records of learning for individuals over time they run the risk of being overwhelmed by the rich data they will have at hand. Without adequate analysis tools to make sense of the data, the benefits to students of better records of learning growth will not be obtained. Processes for managing these records must assume that the records come from wide range of sources (standardised and online tests, observations, class assessments, embedded assessments) and that a graphic history for each student can be displayed.

A range of analytical tools to help teachers understand their data can be anticipated. Issues that will be relevant in developing these tools include:

- Trajectories are idiosyncratic, and may not be able to be projected forward.
- Development of analytical models for individual development analysis is in its early days.
- Based on (possibly dated) research, and a possibly overly constrained learning process (CCC), it was necessary to estimate some individual parameters to project the likely pathways of development. Group data can be used to estimate some parameters only.
- Models based on the previously achieved point and previous estimates of rates of change are the most useful predictors of the next learning status point at t=x (implied in Molenaar et al. 2009 and Malone et al. 1979).
- At any point, the value of the learning scale has meaning in terms of what is says about a student’s likely skills.