

# **School Reform – An exploratory case study of the impact of student centred learning in two primary schools.**

**Chris Goldspink** (corresponding author)

Research Fellow

Inceptlabs, Sydney, Australia

cgoldspink@inceptlabs.com.au

**Margot Foster**

Manager Learning to learn

Department of Education & Children's Services South Australia

Foster.margot@saugov.sa.gov.au

## ***Abstract***

This paper documents a case study into the effect of the adoption of constructivist pedagogy in two primary schools. The cases were chosen due to their similar geographical and socio-economic location and student demographics. One had undertaken considerable pedagogic and site transformation consistent with a South Australia school reform initiative known as Learning to Learn. The other was about to enter the reform program but had not undertaken systematic change at the time of the study. The aim was to trial several methods for measuring the experience of learners in the reformed context and the degree to which this impacted on social and academic outcomes. In addition to providing information on the effectiveness of the measures trialled, the study provides evidence of a significant difference in academic achievement, engagement and meta-cognition resulting from the reform.

**Key words:** School reform, constructivism, outcome measures, case study

**Word count:** 6000

# **School Reform – An exploratory case study of the impact of student centred learning in two primary schools.**

## ***Introduction***

Over the past nine years the Department of Education and Children's Services (DECS) in the State of South Australia has pursued a program to transform the work of educators from one of teaching to that of supporting learning. The change project is called 'Learning to Learn' and has so far involved 180 schools, (33% of all state schools) and thousands of teachers. A program of systematic research has been undertaken alongside the program both to inform the strategy employed and to gauge results. This case study forms a part of that research.

In the past DECS had used only a limited set of academic measures to evaluate performance. These were not regarded as sufficient for evaluating the effect of this change program. A review of the literature turned up few measures which were consistent with constructivist educational practice. In particular, few quantitative instruments were available for measuring variables such as student involvement, wellbeing, pedagogical quality and higher order learning outcomes. The intent of this research was to trial alternative existing approaches or to develop new methods for measuring the effect of change of school culture and teaching practice on student learning outcomes. The aim was not to undertake full-scale development of methods as too little was understood about how best to proceed and what may work, rather it was to gain insight into methodological issues and to identify approaches which had sufficient promise to warrant further investment.

## ***Case Study Design***

The study focused on the following questions:

1. How do children experience learning differently where a student focused approach is integrated into the school culture and the curriculum adopted is consistent with constructivist assumptions about learning?
2. What are the collective outcomes of such change measured using traditional academic measures, meta-learning criteria and indices of student wellbeing?

The cases were two South Australian Primary Schools: One school had undertaken substantial constructivist inspired reform and the other followed traditional practice. Anecdotal evidence as well as existing qualitative and student achievement data indicated that the reform school enjoyed significantly better performance in attendance, retention, discipline, wellbeing and both traditional academic as well as thinking and problem solving outcomes. The reform school in the study will be referred to as River-school Primary and the comparison school is referred to here as Woods-borough.

### **River-school**

- School population: 201 (FTE)
- Years: R-7

- Index of disadvantage: 7 (1 is very disadvantaged, 7 is very advantaged)
- Distance to GPO: 25km
- School Card holders: 51(FTE)
- Attendance Rate: 92.2%

The matching non-reform school is Woods-borough Primary. Woods-borough had only just commenced with Learning to Learn.

**Woods-borough**

- School population: 198 (FTE)
- Years: R-7
- Index of disadvantage: 5
- Distance to GPO: 43km
- School card Holders: 33
- Attendance Rate: 93.3%

Cohorts of students for inclusion in the study were chosen in each school within the following year bands.

- Early Years (Reception to year 2)
- Primary years (years 3-5)
- Middle schooling (years 6-7)

The simplified model below was adopted to guide the work. This implies that quality of pedagogy (and in particular the adoption of constructivist pedagogy) was anticipated to impact on short term indicators of outcomes such as engagement and wellbeing and that this in turn would impact on learning outcomes.

**Figure one: Primary relationships between key variables**



**The instruments**

It was immediately apparent that there were significant gaps in coverage of suitable methods in the area of most concern to this study. In particular: the relationship between pedagogy and student engagement and wellbeing; short term academic; and long term (in particular meta-cognitive) learning outcomes.

To gain information about students’ experience of learning two questionnaires were developed, one for students and one for teachers. The student questionnaire combined both open and closed questions which asked about the students experience at the time of the data collection. This related to what they were doing, where they were doing it and towards what learning aims (Unger 2003). Also included were questions about how much influence students perceived they had over what and how the learning was

being undertaken. They were also asked questions about if they were working alone or with others and how much choice they had in the social arrangements of learning. Finally they were asked if they judged that they had been on task at the time of the survey and if they were interested in what they were doing.

The teacher questionnaire canvassed what the teacher was doing and how they conceived of the role they were performing at the time. The teacher was also asked to rate the level of on-task behaviour evident and to comment on their own level of interest.

## **Measuring Engagement**

The model in figure one makes clear that pedagogy was anticipated to have a short term effect on what we called engagement and that engagement was expected to moderate achievement. The literature review had revealed, however, that the concept of engagement was ill defined and that there were many different ways of measuring it. Most of the established approaches were inconsistent with the needs of this study. We became aware of work being undertaken by Pam Winter (2003) on student involvement and wellbeing in early years. This was derived from prior work undertaken by Laevers (1994; 1997) and Mayr & Ulich (1999). Winter had also developed observational scales for measuring these variables and these appeared suitable for our needs if adjusted to accommodate a wider range of student ages. This modification was undertaken in collaboration with Winter. An additional scale was developed to measure thinking skills. The observation instruments used in the case studies have subsequently been developed through a separate exercise involving a larger and more diverse sample (Goldspink 2008; Goldspink and Winter 2008). This included the development of self-report instruments to complement the observation schedules. For the research reported here however, only the observation scales were available.

## **Measuring Outcomes**

Outcomes data was to include the standard data available from the Departmental standard testing. The Wellbeing measures included in the Engagement concept were regarded as outcomes measures as well as process indicators. With respect to thinking skills, the literature review turned up some relevant work being undertaken within the OECD, driven by the Finnish Learning to Learn (Hautamaki et al. 2002; Hautamaki and Kupiainen 2002). This canvassed similar variables to this study (in particular student background, attitude, wellbeing and thinking skills). English versions of some of their instruments were included in this study for experimental purposes.

## **Time Sample Method**

As preliminary investigation revealed that we were to be comparing schools with very different pedagogy, a methodology was required which would collect valid and reliable data across a spectrum of alternative approaches to teaching and in widely varying contexts. As the aim was to sample actual experience in response to the specific approach to learning within each school, there was a need to collect data which was spread over a wide range of experiences and to collect data as near to the actual experience as possible to avoid problems associated with memory, attribution or rationalisation. The Experience Sampling method was judged the most appropriate (Christensen et al. 2003; Scollon et al. 2003).

The fully randomised time sampling approach commonly used with Experience Sampling was not used as it would have collected a significant amount of irrelevant (i.e. non learning experience related) data. Instead the school day was divided into 3 time periods (excluding breaks). Data was then collected twice a week on random days at, or close to, 10am, 12pm and 2:30pm for each level cohort within each school. At the collection time, both adults and students completed the questionnaire about what they were doing prior to the interruption. Each age cohort within each school was sampled over a period of one month. Spreading the study over a one month period meant that it was conducted during a reasonable sample of a school year and was not subject to short term influences (such as early term periods or special project periods) it also minimised intrusion on the normal operations of the schools.

Immediately prior to the interruption of learning to collect questionnaires for the experience sample, observers chose two of the participating children at random for observation. The observation collected wellbeing data. A teacher questionnaire was also administered at this time – with questionnaires being completed by all teachers or adult staff involved with the cohort at the time of the sample.

The methodology resulted in eight experience samples per cohort. The total number of completed self-report questionnaires collected was therefore 1266. The total number of observation forms returned was 228 while the number of teacher forms was 72.

## **Results**

### ***The Experience of Learning***

As expected the school had a statistically significant association with the pedagogical choice variables of ‘choice in what learn’ and ‘choice in how learn’. While significant statistically they did not however account for a substantial amount of variance. Subsequent analysis of these relationships using regression and structural equation modelling indicated that there was also a direct effect of school – probably due to a difference in school climate on engagement variables. This indicates that these variables were sensitive to whole school as well as class specific factors, particularly the choice variables measured. School in particular was strongly correlated to wellbeing and wellbeing to the depth and quality of thinking observed.

**Figure two: Correlations (Pearsons) between School and Engagement Variables**

	School	Choice in what learn	Choice in how learn	Interest	On-task	Wellbeing	Thinking
School	-						
Choice in what learn	.181*	-					
Choice in how learn	.134*	.467**	-				
Interest	.382**	.319**	.294**	-			
On-task	.282**	.159*	.225**		-		
Wellbeing	.462**	.095	.111	.225**		-	
Thinking	.242**	.014	.029	.092	.077	.630**	-

\*  $p < 0.05$ . \*\* $p < 0.01$ .

Examining these relationships in more detail revealed that students at River-school reported having had ‘a lot of choice’ on 52% of occasions compared to a little more than half this rate at Woods-borough (27%). These figures were very similar with respect to choice of *how* to learn. Students at Woods-borough Primary reported having ‘no choice’ in *how* they were learning in 45% of cases, compared to less than half this rate (20%) for River-school. All these differences were statistically significant ( $p < .0001$ ).

Students at River school enjoyed a wider range of teaching methods. At the time they were stopped to complete the questionnaire, the most common response by students in River-school about *how* they were learning was ‘listening to a teacher or adult’ (20% of the time). This compared to 30% of the time for Woods-borough. The second most common report was ‘thinking by self’ (19.2%). This was comparable with the figure reported for Woods-borough (22%). River-school students were almost twice as likely to report learning by talking with other students (11% compared with 6%), and three times more likely to be producing information using e-media (6% compared to 2%) or to be using physical activity or making things (7% compared to 2%). Students at both schools were approximately equally likely to be learning by writing (12% compared to 13%) and Woods-borough students were slightly more likely to be learning by reading (9% compared to 7%).

River-school students were working at their desk on 34% of the occasions surveyed compared to 63% of cases for Woods-borough: a further 56% of the time they reported being elsewhere in the class. Woods-borough students were a little more likely to be away from the class altogether (14% compared to 10% for River-school).

## Engagement

### Interest and on task behaviour as reported by students

Overall students in both schools indicated being ‘very interested’ on 46% of the occasions they were asked. A further 32% indicated that they were ‘a little interested’. Some 4.2% said they were ‘bored’ and a further 9% ‘very bored’. However, these figures varied markedly by school. At River-school, students identified being interested to some extent on 88% of occasions. This was comprised of 55% who said ‘very interested’ and a further 33% who said ‘a little interested’. For Woods-borough,

students indicated being interested on 62% of occasions, comprised of 33% 'very interested' and 29% 'a little interested'. Correspondingly, in River-school students identified being bored (2.9%) or very bored (2%) on 5% of occasions compared to Woods-borough where students reported 'bored' (6%) and 'very bored' (20%) on 26% of occasions. All of these differences were statistically significant (  $p < .0001$  level 2 tailed).

There was a small but statistically significant correlation between 'where working' and level of interest ( $r=0.167$ ,  $p=.01$  2 tailed). Sixty four percent of those who said they were 'bored' and 60% of those who said 'very bored' were working at their desk compared to 34% of those who said that they were very interested.

Overall some 45% of all students sampled at River-school indicated that at the time the data was collected they were 'on task' to a 'very high level'. This compared to 19% for Woods-borough. There were no students at River-school that reported low levels of on-task behaviour.

### **Observed Wellbeing.**

The following data was collected using modified versions of the observation instruments developed by Winter (2003). These included three scales: happiness and satisfaction, social functioning and dispositions to learning, which combined to form a single composite scale of wellbeing. The data collected by observation comprised a sub-sample of the students who provided self-report on their experience of learning.

River-school had the highest level of happiness and satisfaction with 77% having high compared to Woods-borough with 66%. No students were rated as having 'low' Happiness and Satisfaction in River-school, while 7% of those at Woods-borough were so rated. This led to a mean score of 2.77 (s.d .425) (on a three interval scale) for the reform school and 2.6 (s.d. .625) for the non-reform. The standard deviations show that as well as having a higher average the reform school had less variance.

The school based assessments varied markedly for Social Functioning. Sixty five percent of River-school students observed were rated as demonstrating 'high' social functioning. This compared with 43% for Woods-borough. A further 34% of River-school students were rated as having medium social functioning compared with and 50% for Woods-borough. Only 1 child was rated as having low social functioning when observed at River-school compared with 8 at Woods-borough. The mean score for the reform school was 2.63 (s.d. .509) compared to 2.36 (s.d. .616) for the non-reform school.

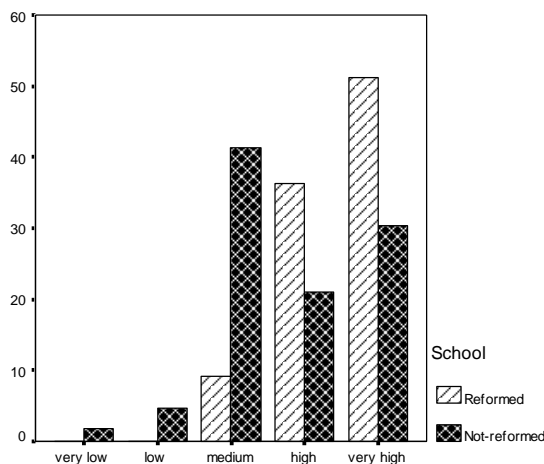
The significant difference in rating by school was also reflected in assessment of students' Dispositions to learning. Seventy four percent of River-school students observed were rated as 'high' on this scale, this compared to 48% of those at Woods-borough. Twenty six percent of River-school were rated as having a 'medium' score on this scale compared to 47% of Woods-borough. Woods-borough was the only school to have students in the low category with 12 (5%) so observed. The mean score for the reform school was 2.74 (s.d. .442 ) compared to 2.27 (s.d. .655) for the non-reform school.

Seventy percent of River-school students observed were rated as demonstrating a high level of thinking compared with 54% of Woods-borough. 39% of Woods-borough demonstrated medium level thinking compared to 30% for River-school. Again Woods-borough was the only school to have students rated in the low category with 8

(3.7%) so rated. The mean score for the reform school on this variable was 2.7 (s.d. .462) and for the non-reform 2.46 (s.d. .637).

River-school scored most highly overall on Wellbeing. Ninety one percent of all students observed were rated either as having ‘very high’ (53%) or ‘high’ (38%) levels of wellbeing. By comparison, only a small majority of Woods-borough students (51%) were overall positive with 30% being rated very high and a further 21% high. A further 42% of Woods-borough students rated as medium and 2% as low. The mean for the reform school (using a five point scale for this variable) was 4.44 (s.d. .663) while that for the non-reform was 3.74 (s.d. 1.01).

**Figure three: Overall Wellbeing by School**



### ***The relationship between Pedagogy and Engagement***

The ‘interest’ variable included in the student questionnaire is a proxy for the individual’s disposition or motivation while the ‘on task’ variable is a proxy for the behaviour of involvement. Measuring the relationships between the ‘level of interest’ and ‘on task’ variables revealed a correlation of  $r=.569$  ( $p<.0001$ ) – the two variables are highly correlated.

Overall, students identified as being ‘on task’ to a ‘very high level’ on 43% of occasions and ‘high’ on a further 30% of occasions: they were on task on nearly three quarters (73%) of all occasions they were asked. On the negative side, students reported being ‘on task’ to a ‘low’ degree on just 3% of occasions and ‘very low’ on a further 5%. There were significant differences by School. River-school students reported ‘very high’ (48%) or ‘high’ (31%) on a total of 79% of occasions compared to Woods-borough students reporting ‘very high’ (34%) or ‘high’ (28%) on a total of 62% of occasions. On the opposite pole, River-school students reported only 3% ‘low’ or ‘very low’, compared to 16% for Woods-borough. All of these differences were statistically significant. There was also a statistically significant difference by gender but the difference was small.

Turning now to indicators of a student centred learning approach – the *choice* variables: a strong correlation ( $r=.404$ ,  $p<.0001$ ) was measured between student reports of ‘how interested were you’ and how much choice they reported in *what* they were learning. As this is only a correlation it cannot be assumed that the relationship is direct. However, what we can say is that in an environment where students enjoy greater choices in what they learn there appear to co-exist variables which,



individually or in concert, contribute to enhanced student interest and, as interest and 'on task' behaviour are correlated, that choice is associated with higher engagement in task. The small number of observations precluded any detailed analysis of the relationship between pedagogy variables such as choice and the wellbeing variable.

## ***Learning Outcomes***

While there have been a number of tests derived to assess student problem solving and meta-cognitive skill. Almost all of these are designed to be applied to domain specific skills – eg maths, English or science. There are few approaches designed to elicit non-domain specific ability. In part this reflects a debate and general disagreement about the degree to which such skills can be separated from specific domains of learning. For this study we decided to trial three means for examining general thinking skills. The first comprised a score for meta-cognitive level derived from the student's responses to three open questions in the questionnaire. The second was the observer assessment of thinking skills. The third were the results of student's response to a problem solving task drawn from the Finnish Learning to Learn inventories.

### **Meta-cognitive level**

The most experimental of the measures we adopted was that for meta-cognitive level. In the student questionnaire respondents were asked to write in their own words a descriptive answer to the following three questions.

- Tell us what you were just doing
- Why were you doing what you were doing?
- What were you learning?

The responses to these questions were coded to produce a single interval variable. The coding procedure was as follows:

In response to Question one 'Tell us what you were just doing'

- Score 0 if no response is provided (field is blank)
- Score 1 point if a simple description of a task is provided.

In response to Question 2 'Why were you doing what you were doing?'

- Score 0 points for passive direction following e.g. 'the teacher told me to'
- Score 1 point if the reason given for doing it linked to a passive learning goal e.g. 'needed to learn about fractions'
- score two points, if linked to an active learning goal e.g. 'I wanted to find out about how x worked'

If in response to the question 'What were you learning?'

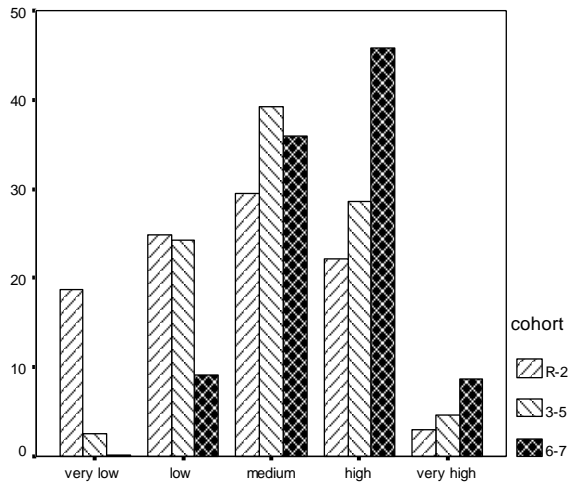
- Score one point if the respondent identified a particular skill or linked the learning to a wider knowledge base.
- Score two points if they identified the learning as a part of a learning strategy or identified a point of reflection on a personal learning need or goal.

The minimum score is 0 and the maximum is 5.

The resulting meta-cognitive level variable provides an indication of the student's ability to reflect on his/her learning and to use the language of learning to describe it. It is a proxy indicator of their thinking ability and ability to consider and possibly take responsibility for managing their own learning. We expected that it would be sensitive also to the students literacy level.

The mean meta-cognitive level was 2.16 and the Standard Deviation was .965. Not surprisingly meta-cognitive level varied by age cohort. The age distribution is shown in the following chart.

**Figure four: Meta-cognitive level by age cohort**



This variation was not however to the degree as may be expected. At the low end it was unsurprising to find that of all those who scored very low meta-cognitive level, 77% were in years R-2. However, given that the coding criteria were applied the same way to all students, it is perhaps surprising to find that of those who scored very highly, 11% were in years band R-2, 33% in 3-5 and 56% in year 6-7.

The following table summarises the relationships between the school, engagement variables and outcome measures.

**Figure five: Correlations (Pearsons) between Engagement and Outcome Variables**

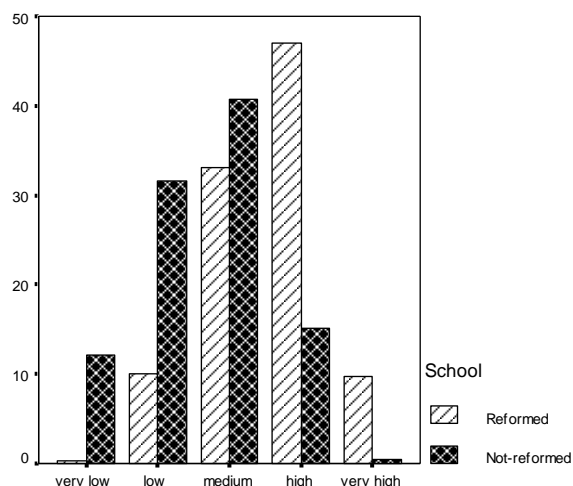
	School	Wellbeing	Thinking	Interest	On-task	Standard Academic score	Meta-cognitive level	Mean correct on problem solving
<b>School</b>	-							
<b>Wellbeing</b>	.462**	-						
<b>Thinking</b>	.242**	.630**	-					
<b>Interest</b>	.382**	.225**	.092	-				
<b>On-task</b>	.282**	.155	.077		-			
<b>Standard score</b>	.300**	.301**	.300**	.004	.025	-		
<b>Meta-cognitive</b>	.444**	.183*	.186*	.233**	.162*	.228**	-	
<b>Mean correct</b>	.455**	.494**	.285	.215	.105	.259	.551**	-

\* p < 0.05. \*\*p < 0.01. Note that N varies for different scores and is lowest for mean-correct due to this task only being undertaken by the highest age cohort.

The standard score is an average of the score attained on the Departments standard numeracy and literacy tests: it provides a measure of academic achievement. The meta-cognitive index was described above. The mean-correct score is the mean number of questions scored correctly on the Finnish problem solving and logic task.

Again looking more closely at these relationships a significant difference was found between schools. As is clearly evident in the following chart, River-school scored most strongly on the meta-cognitive index with only two children (0.3%) scoring very low, 10% low, 33% medium 47% high and 10% very high. The between site difference was significant (p< .0001 level 2 tailed).

**Figure six: meta-cognitive level by school**

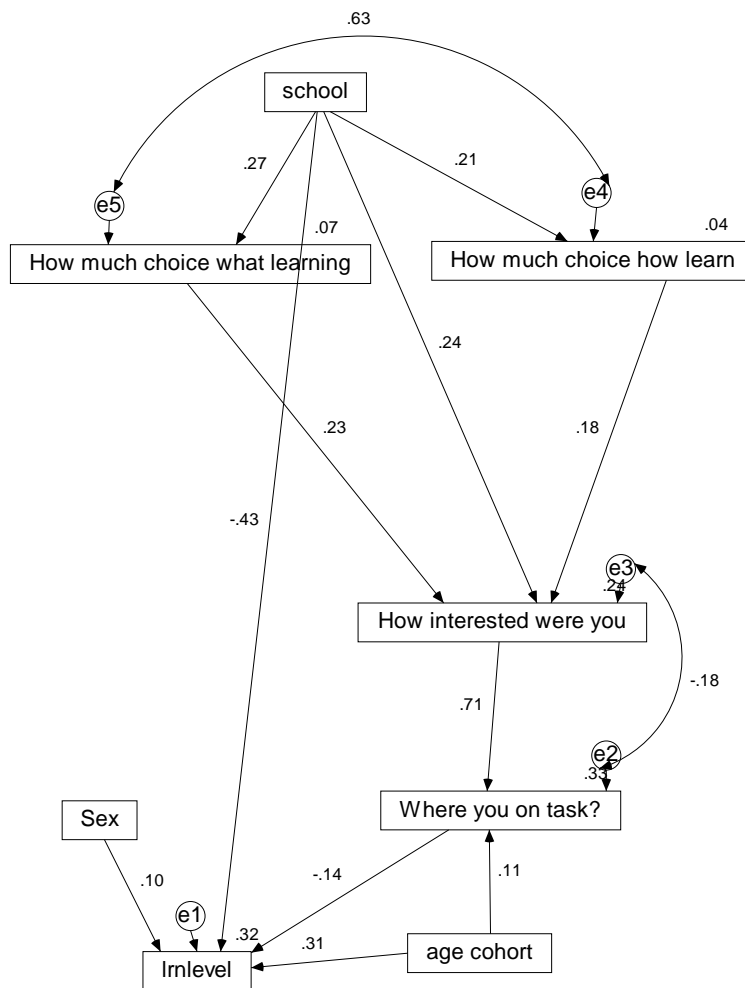


The difference in meta-cognitive score varied significantly by gender. This was apparent across all age cohorts but was most marked in the R-2 group. In this year band 27% of males scored very low – nearly three times the rate for girls (9.7%), a further 30% of boys scored low compared to 20% for girls. As the assessment was based on the capacity to express their experience in words, this may reflect more

limited literacy in the male cohort compared to the female. The gender difference proved statistically significant once age was controlled for.

The following structure diagram was produced using the software package AMOS 7. Maximum likelihood estimation was used. For analysis using structural equation modelling the position of Hu and Bentler (1999) is adopted with respect to determination of fit. It is generally not expected that the Chi square statistic (CMIN) will prove insignificant ( $p < 0.05$ ) rather the less stringent CFI and RMSEA statistics are used. For the model to be accepted as in fit the CFI would need to be  $> 0.95$  and/or the RMSEA  $< 0.05$ . Unless otherwise indicated standardised measures are presented. All analysis is performed on the covariance matrix.

**Figure seven: Structural analysis of relationships effecting meta-cognitive level.**



CMIN 102.225,  $df=15$ ,  $p=.000$ ; CFI=.961; RMSEA=0.068

The diagram combines all of the variables anticipated to have an effect on the Meta-cognitive levels of students. The model is in fit according to the CFI measure. All included relationships were statistically significant. From this diagram it is apparent that school predicts both degree of choice of *what* and *how* to learn with school accounting for an estimated 7% of variance on choice *what* and 4% of variance of choice *how* respectively. Note the very large co-variance between error terms e4 and e5. This suggests that there is one or more variable, independent of school, which influences degree of choice. There are several possible candidates: teacher style and student perception prominent among them. We expect that teacher style would be a

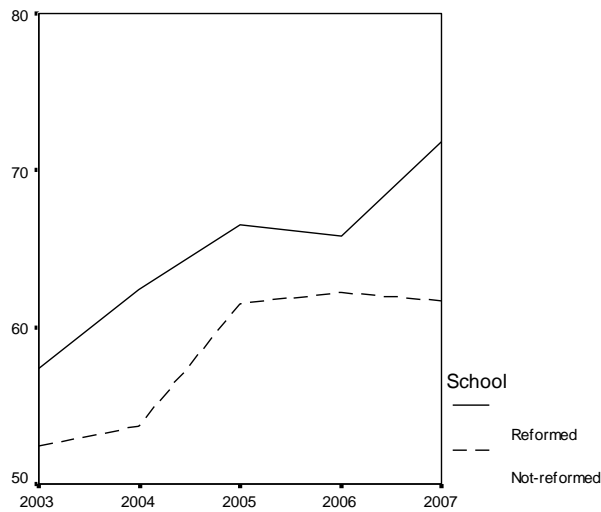
significant contributor. Progressing down the diagram, it is apparent that the degree of choice of *what* learning predicts level of interest accounting with an effect size of  $r=.23$  while choice in *how* learn also predicts interest but to a lesser ( $r=.18$ ) degree. The direct and indirect effect of school on interest accounts for an estimated 24% of variance. The relative difference between choice in *what* and *how* is quite sensible when it is appreciated that interest has been found to be subject specific (Schiefele 1991; Hidi and Renninger 2006) meaning that choice in *what* learn should be expected to make a greater difference to personal interest. Note however that the largest effect size ( $r=.24$ ) on interest is the school itself. This indicates that there is a difference made by school which is not associated with choice in what and/or how to learn. This is most likely attributable to wider cultural aspects of the school which impact on situational interest. The level of interest predicts 'on task' to a very large degree ( $r=.71$ ) accounting for an estimated 33% of variance. There is some residual error which is correlated however – evident in the covariance of 0.18 between the error terms  $e_2$  and  $e_3$ . Again this suggests an additional variable which is not in the model. This is likely to be either something associated with teaching (quality pedagogy can generate involvement even in the absence of individual and situational interest by means of extrinsic motivation) or the individual student (short term distractions, difficulties in concentrating etc) or both. As was to be expected, age predicts meta-cognitive level so too does gender. The largest contributor however is school with an effect size of  $r=0.43$ . There is a further indirect effect of school via the pathway of *interest* and *on task*. Together these variables explain an estimated 32% of variance on meta-cognitive level. All of these relationships were statistically significant ( $p<.001$  two tailed).

## **Academic Achievement**

Standard Departmental numeracy and literacy test results were obtained for as many of the students that participated in the study as possible. These tests are not collected every year so not all students were covered by results for the year of the study. They were also only available for the 3-5 and 6-7 age cohorts. The tests address both literacy and numeracy and the results are not age relative. For all of the following analysis the Standard score variable produced from the test results was used as it presents a continuous scale variable most suited for use in regression and SEM.

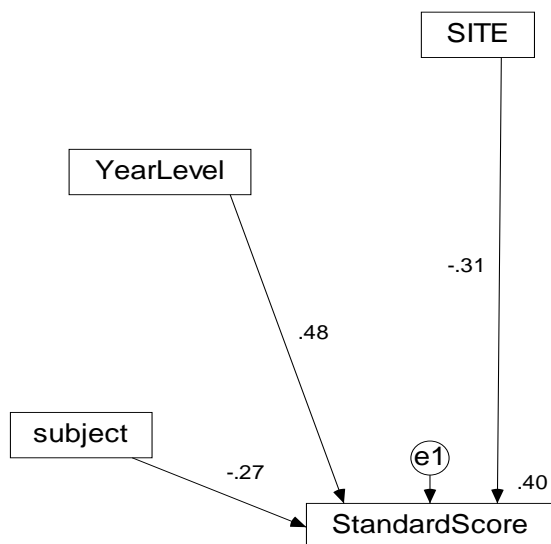
As with the other outcome variables, the reform school demonstrated consistently better results than the non-reform school in both literacy and numeracy.

**Figure eight: Standard score (numeracy and literacy combined) by year by site**



The following structure diagram shows that a total of 40% of variance on the standard score for the tests is explained by the three variables, year band, school and subject. As the standard score is not age relative it was to be expected that year level would have a significant effect. This is supported in the results by a standardized regression weight of  $r=.48$  meaning that this variable alone explained 23% of total variance. The next most significant effect is from the school with a standardized regression weight of  $r=.31$ : school differences therefore explained 9.6% of total variance. The subject variable captures the difference between numeracy and literacy scores and explains an additional 7.3% of variance. This model proved to have an excellent fit using all three fit indexes meaning that it closely reflects the relationship evident in the data.

**Figure nine: variables influencing LAN scores.**



CMIN 3.1,  $df=3$ ,  $p=0.375$ ; CFI=0.997; RMSEA=0.02

Taking the data for all years for which data was available for the students included in the study (i.e. not just the year of the study but from 2003 to 2007) the effect of differences between the two case schools was examined using a series of regression

models. The first set of models was run on the combined standard scores for numeracy and literacy across all years for which data was obtained.

Overall some 35% of total variance was explained by the final model which included both year-level and site. As is to be expected, the largest proportion of variance (27.7%) was explained by year-level. The site explained a further 7.9% ( $R^2=.079$ ,  $p=.000$ ). In other words approximately 8% of the difference in the scores can be attributed to differences between these schools. As the case schools were chosen to be similar in all respects except culture and pedagogy, differences achieved through the changes made in these areas should account for the main component of this difference. Note that this is significantly larger than the difference found in similar studies which focus on the relationship between pedagogy and student academic achievement (see for example Doherty and Hilberg 2007).

A gender specific effect was noted for the literacy data. A further series of models was run on the literacy result for all year data obtained to provide an estimate of the effect size for this variable. For the literacy results, a total of 49.6% of variance was explained by the combination of year-level, gender and school. Again year-level explained the greatest proportion of this (35.8%), gender a statistically significant 5.8% and site a further 9%.

From these results we can conclude that all other things being equal (i.e. children's intrinsic ability, background of parents, socio-economic influenced factors in community etc), differences in the two schools account for a statistically significant proportion of variance on academic achievement scores in both numeracy and literacy. School differences appear to account for approximately 9% of variance – or one third of the effect size of year-level. This is a very substantial effect. Care needs to be taken however in generalising this due to possible hidden differences in the schools and the fact that the chosen schools, while similar to one another, are clearly not generally representative.

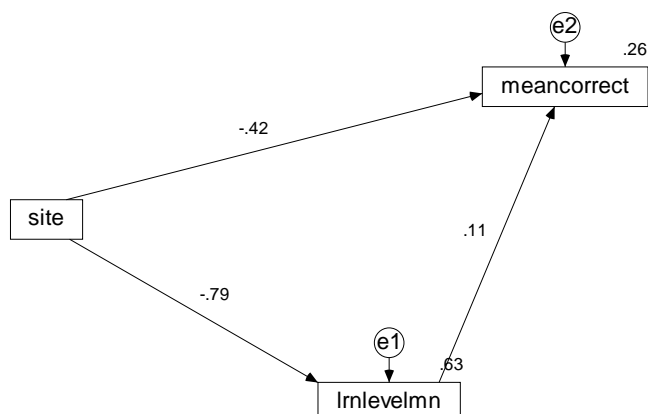
## **The Finnish Inventory**

As a part of this study, several inventories from the Finnish Learning to Learn project were trailed. While still in development, these have now been used with large samples in a number of OECD countries. The instruments are demanding and time consuming to complete and only one of the meta-cognitive exercises was completed by both case schools. The test undertaken by both case schools comprised thirteen questions.

Examining the simple scores, River-school had a higher proportion of students with the correct answer in ten out of the thirteen questions. The relationship of these answers were compared to learning level and site using a mean correct answer score calculated by summing the correct scores for all students within a site and dividing by the number of students at that site. The first significant finding is that the meta-cognitive level derived from student answers as discussed above was significantly correlated to the mean-correct score ( $r=.409$ ,  $p<.01$ ). Both are designed to measure meta-cognitive thinking skills. This correlation gives some confidence that they are indeed measuring a similar underlying construct. As with the meta-cognitive level score, mean-correct was strongly correlated ( $r=.508$ ,  $p < .01$ ) with school.

To further test the relationship between site, the meta-cognitive index and the performance of students on the thinking skills task, the following structural equation model was tested.

**Figure ten: Structural model of the relationship between site, meta-cognition and the thinking task score**



Note that this model has zero degree of freedom so the fit indices are meaningless and are not-reported. What is significant about this model (and a little disappointing as we were hoping that the meta-cognitive index may be an adequate proxy for thinking skills) is that while site predicts the meta-cognitive score to a very high degree ( $r=.79$ ) and site also predicts the mean correct score on the thinking test ( $r=.42$ ), there is only a marginally significant relationship between the meta-cognitive score and thinking skill performance. What this means is that while the difference in the schools explains a very substantial amount of both capacity to articulate about learning (63% of variance on meta-cognitive index) and the capacity to perform well in a problem solving task (the two paths together explaining an estimated 26% of variance on the test performance) the meta-cognitive index does not strongly predict performance on the task. In other words, student's who cannot articulate their thinking can nevertheless perform well on the problem solving task and that students who can articulate their thinking may not perform well on the problem solving task. This result needs to be interpreted with some care. Firstly, the problem solving task used was only one of a number developed by the Finnish learning to Learn project and so it is possible that a difference would be more prominent on tests other than the one trailed. In addition, the idea that higher order thinking can be considered independently of a domain of study (maths, language, science) remains controversial and so therefore must the veracity of these tests overall. Clearly further work is needed in this area.

## Findings

The school which had undertaken considerable cultural transformation and pedagogical change in a manner consistent with constructivist principles of learning was River-school Primary. Evidence of the change was reflected in the differences students experienced in learning. In particular:

- greater choice in what to learn;
- greater choice in how to learn;
- learning was more social;
- learning was less class centred;
- teachers' were more interested.



These differences, as well as some unmeasured difference in site (possibly a culture of regard), contributed to statistically significant:

- higher 'on task' behaviour associated with greater interest and less boredom.
- high happiness and satisfaction
- higher social functioning
- markedly higher disposition to learning and
- higher observed thinking skills
- markedly better overall wellbeing
- better problem solving ability on the assigned task
- higher capability to articulate about learning (meta-cognition); and
- superior performance on standard numeracy and literacy tests

## **Conclusions**

On the basis of the experience gained through this trial we can conclude that it is possible to measure the effect of change associated with the adoption of constructivist pedagogy and to trace its effect through short term engagement and wellbeing and on learning outcomes. In this study, statistically significant relationships were established using relatively simple instrumentation and a modest sample. The following observations are made with respect to their performance:

While student choice is a goal of and an indicator of pedagogical practice associated with a transformed school, choice alone does not capture the full extent of the pedagogical experience and therefore does not fully account for the measured difference on immediate scores of engagement and wellbeing or on higher order outcomes. More sophisticated measures of pedagogical and site cultural quality are desirable.

The observation of wellbeing is feasible, however, it is somewhat time consuming (and hence expensive) and is subject to a variety of biases. Wellbeing would be better measured by self-report where this is feasible (i.e. with older age groups). Where observation is used, well trained and site-neutral observers should be used. Ideally observation measures should be triangulated with either self-assessment scores or teacher' scores or both.

Meta-cognitive skill can be gauged based on student responses to three simple open ended questions. This data, while easy to capture, is time consuming to code. Its relationship to the ability to successfully complete problem solving tasks requiring higher order thinking skills is however questionable and needs to be further examined. It does appear feasible to use problem solving tasks (like those being advanced by the Finnish Learning to Learn team) to assess thinking skill independently from a particular subject. The tasks are themselves quite complicated and take time to administer and present quite complicated data entry requirements. To make them widely acceptable it may be advisable to turn a set into a game which can be played on-line and with data captured automatically as the game is played.

The results of the case study confirmed the anecdotal evidence that the adoption of constructivist practice can have a positive effect on students experience of learning

and that this can in turn lead to improved social, meta-cognitive and academic outcomes. The findings of the effect size are quite large and, while caution is needed in generalising from them, they justify further investment in systematic research, including research which can better evaluate the effect of student background variables and alternative school contexts and environments. A larger sample and a wider spread of schools would be highly desirable.

At the time of writing plans are in place to further test these measures with a sample of some 20 schools involving a wide range of subjects and involving approximately 1000 students.

## References

- Christensen, T. C., et al. (2003). "A Practical Guide to Experience Sampling Procedures." Journal of Happiness Studies **4**: 53-78.
- Doherty, R. W. and R. S. Hilberg (2007). "Standards for Effective Pedagogy, Classroom Orientation, English Proficiency , and Student Achievement." The Journal of Educational Research **101**(1): 24-34.
- Goldspink, C. (2008). Towards a framework and instruments for the measurement of the effect of pedagogy of student wellbeing. Social Futures Institute conference "Health, Wellbeing and Happiness: from Local Action to Global Change". University of Teesside, Middlesbrough, UK.
- Goldspink, C. and P. Winter (2008). Student Engagement and Quality Pedagogy. European Education Research conference. Gotebourg, Sweden.
- Hautamaki, J., et al. (2002). Assessing Learning to Learn: A Framework. Helsinki, Centre for Educational Assessment.
- Hautamaki, J. and S. Kupiainen (2002). The Finnish Learning-to-learn Assessment Project: A Concise Report with Key Results. Brussels, Centre for Educational Assessment.
- Hidi, S. and K. A. Renninger (2006). "The Four-phase Model of Interest Development." Educational Psychologist **41**(2): 111-127.
- Hu, L.-t. and P. Bentler (1999). "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives." Structural Equation Modelling **6**(1): 1-55.
- Laevers, F. (1994). Defining and assessing quality in Early Childhood education. Studia Paedagogica. Leuven Leuven University Press.
- Laevers, F. V., E.; Kog, M.; Depondt, L. (1997). A process-oriented child monitoring system for young children. Leuven, Centre for Experiential Education, Katholieke Universiteit Leuven.
- Mayr, T. and M. Ulich (1999). "Well-being in Day Care Centres: an exploratory empirical study." International Journal of Early Years Education, **7**(3): 10.
- Schiefele, U. (1991). "Interest, Learning, and motivation." Educational Psychologist **26**(3-4): 299-323.
- Scollon, C. N., et al. (2003). "Experience Sampling: Promises and Pitfalls, Strengths and Weaknesses." Journal of Happiness Studies **4**: 5-34.

Unger , C. (2003). "Why Listen to Students?" from <http://www.newhorizons.org/voices/unger.htm>.

Winter, P. (2003). Curriculum for babies and toddlers: a critical evaluation of the first phase (birth to age three) of the South Australian Curriculum, Standards and Accountability Framework in selected childcare centres in South Australia. Education. Magil, University of South Australia. **PhD**.