

A Theory-driven and Evidence-based Approach to School Improvement: The Independent Schools Victoria *L E A D* Surveys

Panayiotis Antoniou¹, Jacqui Louw² & Peter Gronn¹

¹University of Cambridge, Faculty of Education, UK

² Independent Schools Victoria, Australia

Abstract

Research evidence suggests that schools should develop evaluation mechanisms and make use of the data emerging from such evaluations to improve teaching practice and the school learning environment. However, it is not always easy for schools to decide on the appropriate self-evaluation framework or identify validated research data collection tools. Such a framework and tools have been developed by Independent Schools Victoria in Australia. *The L E A D Project* is a comprehensive suite of school stakeholder surveys and data analysis, providing benchmarked performance measures for school boards and senior leadership teams. This paper aims to present the theoretical framework of the *L E A D* Suite of School Stakeholder Surveys carried out by Independent Schools Victoria. In addition, the paper aims to explore and present findings relating to the evaluation of the content and construct validity, the reliability and the internal consistency of the *L E A D* parent, student and staff questionnaires. Finally, based on the conclusions of this evaluation, suggestions to schools aiming to engage in a theory-driven and evidence-based approach towards improving their school effectiveness are provided. Data were gathered from a total of 119,749 students, teaching staff, general (non-teaching) staff, parents and school board members taking the *L E A D* Surveys in 112 independent schools in Victoria, Australia, and followed a five-year longitudinal design from 2009 to 2013. Implications of the findings for school improvement utilising a theory-driven and evidence-based approach are discussed and suggestions for further research are also provided.

1. Literature Review

In recent years, national policies have been adopted in many educational systems to promote educational accountability; the quality of the curriculum; the selection of teachers; the provision of teacher initial and in-service training; and the use of resources to improve schooling (Desimone, Smith, Hayes, & Frisvold, 2005). Educational policy at the national level is regarded as an important component of the effort to turn around low-performing schools and enhance student outcomes. For example, the United States educational system employs a top-down accountability model in which schools act as the distributors of national education policies. Individual schools are held responsible for the quality of education they offer to their students, and are required to “provide information to policy makers and the public about value for money, compliance with standards and regulation and quality of the services provided” (Rosenkvist, 2010, p. 8). In this context, policy-makers emphasize centrally developed policies in which school improvement results from macro-level system planning (Levin, 2010). However, one of the tensions which runs through policy development and implementation is between the need to be aware of the general patterns and apparent commonalities or convergence between localities while considering the local and school particularities (Whitty & Edwards, 1998).

Taking this tension into consideration, researchers and policy-makers have paid increasing attention to school self-evaluation procedures and the subsequent development of policy at the school level. A consistent argument among many scholars has been that responsibility for, and control of reform efforts should be located at the individual school level. Schools are seen as the “basic unit of change and school educators (teachers and principals) are not only the agents, but also the initiators, designers, and directors of change efforts” (Smith & O’Day, 1991, p. 235). Along the same line of argument, Spillane (2005) argues that local school systems are more than mere implementers of top-down educational policies. He supports the view that schools should be allowed to respond to national policy initiatives by developing school self-evaluation mechanisms and adopting their own distinct policies. As Flessa (2012) argues, the literature on policy implementation and on schools as organisations tends to view schools as idiosyncratic places that are more likely to change a reform than be changed by it, using site-level autonomy and discretion to redirect policy goals in unexpected ways. Moreover, it is claimed that school decentralisation initiatives increase the flexibility of the schools and allow them to develop school distinct policies which have a better potential to raise student achievement (Caldwell, 2003; Robertson, Wohlstetter, Albers, & Mohrman, 1995). The main assertion is that increasing schools’ authority and flexibility will allow for the development of effective educational processes which are likely to correspond better to local needs. School stakeholders are more aware of their school needs than are policy makers, and may therefore be more efficient at directing effort, resources and educational processes to meet their needs (Nir & Ben Ami, 2006). School stakeholders represent key forces in school policy development (Broadhead et al., 1996; Canner, 1985; Constable, 1994) and they are considered as a lever which is especially well positioned to develop appropriate policies and actions to remedy school underperformance (Vitaska, 2008).

Central to the development of educational policy at the school level is the requisite to identify needs and priorities for improvement in each school. This could be achieved through the adoption and implementation of a school self-evaluation framework, which stresses a school’s own responsibility for quality (Hofman et al., 2010; Creemers, Antoniou & Kyriakides, 2013a;2013b). This is important because, according to Barber (1996), the essence of a successful organisation is the search for improvement and effective self-evaluation is the key to it. School Self-Evaluation (SSE) could be described as a constant quest for evidence in a school’s transparent sense of purpose, behaviour, relationships and classroom performance. SSE is a collaborative, reflective process of school review. Devos (1998) argues that SSE should be seen as “a process mainly initiated by the school to collect systematic information about the school’s functioning, to analyse and judge this information regarding the quality of the school’s education and to make decisions that provide recommendations” (pp. 1–2). It provides school teachers with a means of systematically looking at how they teach and how pupils learn and helps schools improve outcomes for learners.

As Kyriakides and Campbell (2004) argue, the overarching goals for SSE are twofold: To improve the quality of the organisation and to improve teaching and learning. Improving the quality of the organisation includes matters such as social relations between members of the organisation; organisational climate and culture (Freiberg, 1999); the nature of decision-making (Hoy & Miskel, 2001); and the responsiveness of the school as an organisation to external and internal change forces (Fullan, 2001). Improving teaching and

learning involves concepts of teacher effectiveness (Muijs & Reynolds, 2001; Creemers & Kyriakides, 2006); school effectiveness (Creemers & Kyriakides, 2012; Scheerens & Bosker, 1997; Teddlie & Reynolds, 2000); and decisions about how these concepts are appropriately measured within the school setting (Goldstein, 1995). Research indicates that meaningful SSE, focused on teaching and learning and on improving outcomes for pupils, brings about improvement (Creemers & Kyriakides, 2012; MacBeath, 1999). Despite the importance of SSE for school improvement, it is not always easy for schools and school boards to decide on the appropriate and relevant self-evaluation framework or identify validated research tools to collect and analyse data from different aspects of school functioning. In addition, relying only on school stakeholders' experiences can sometimes be limiting to school development and improvement (Britzman, 1991). There appears to be little evidence from published literature that assisting school stakeholders to engage in any improvement program, without providing them with a validated framework and appropriate tools to identify priorities for improvement, necessarily meets their needs or improves school effectiveness (Smith & Hatton, 1992). Such a framework and tools towards school self-evaluation has been developed by *Independent Schools Victoria* in Australia.

2. The Theoretical Framework of the Study – The *L E A D* Surveys

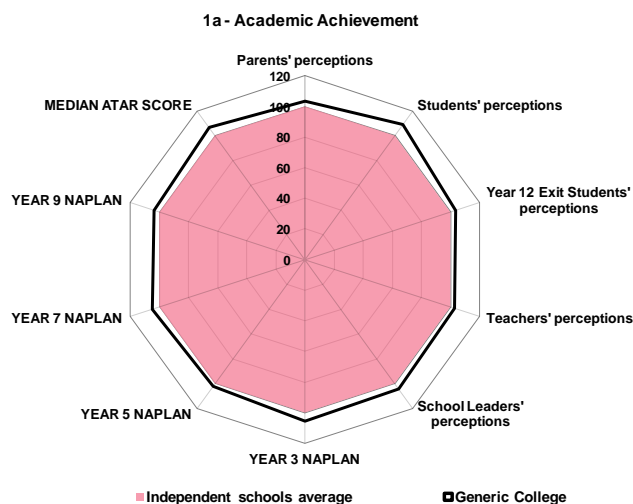
The L E A D Project – 'Listen to our stakeholders, Evaluate what is said, Act on this knowledge and Deliver better outcomes' – is a comprehensive suite of school stakeholder surveys and data analyses, providing longitudinal and benchmarked performance measures for school boards and senior leadership teams.

The *L E A D* Suite of School Stakeholder Surveys are diagnostic and reporting tools that enable school leadership teams to analyse their school's performance against key school effectiveness indicators. School performance is also benchmarked against the results for all participating schools. There are five surveys in the *L E A D* suite of surveys: (1) Parent Satisfaction Survey; (2) Staff Satisfaction Survey; (3) Student Satisfaction Survey; (4) Year 12 Exit Student Satisfaction Survey for students in their final school year; and (5) the Governance (School Board) Survey (which is not discussed in this paper).

In addition, there is a sixth report called *The L E A D Report* which brings together all data sets from the Parent, Student, Staff, Year 12 Exit Student and Governance Satisfaction Surveys into one report. It provides feedback about performance from each of the key stakeholder groups, and aligns these perceptions with real performance data, such as school financial data; teacher salaries; national student test scores (e.g. National Assessment Program – Literacy and Numeracy (NAPLAN) scores for students in years 3, 5, 7 and 9 and Australian Tertiary Admissions Rankings (ATAR) for students completing year 12; student attendance and teacher retention rates; and post-school destinations collated from reliable sources. *The L E A D Report*, which has its roots in the Harvard Business School 'balanced scorecard' approach, was the driving force behind this suite of stakeholder surveys. It is the final stage of each school's involvement in the process. *The L E A D Report* is considered to be the most unique feature of the survey suite, and provides those responsible for school performance with a helicopter view of the complete data set. Each element of the report is underpinned by a richer data set that can be used for additional analyses.

Figure 1 below provides an example of how the data are brought together for each domain of school effectiveness using a radar chart. The sub-domain shown here is 'Academic Achievement'. The stakeholders' perceptions (subjective data) of the school's academic achievement are presented in lower case (parents, students, year 12 students, teachers and school leadership). Objective data (annual national test results for the school – the NAPLAN and ATAR scores) are presented in UPPER CASE. In each radar chart, data for each school are presented as the bold black radar line against the coloured background (the benchmarks). For each indicator on the radar chart, the benchmarks and the school's results have been averaged to a score of 100. The school's score is calculated as a relative score to the benchmark to enable a straightforward comparison. If a school's results extend beyond the benchmark (the coloured background), this can be read as achieving above the average. If a school's results fall short of the benchmark, this can be read as achieving below the average. The charts therefore help schools to identify to what extent there are differences in each stakeholder group's perceptions, as well as how well their perceptions match up to the measured objective indicators.

Figure 1: Example of the data presentation in *The L E A D Report*



2.1 Development of the Parent Satisfaction Survey

The *L E A D* School Stakeholder Surveys were developed by Independent Schools Victoria over a few years, beginning with a pilot parent satisfaction survey in 2006. The aim was to learn what parents thought about the quality of their child's educational experience at their school in order to help identify areas for school improvement. At the time, there was little research measuring parent satisfaction with schooling. King & Bond (2003) reported on the development of a 20-item scale to measure parent satisfaction in Queensland government schools. This was a global measure, although no evidence was presented as to the possibility of coherent subgroups of items being present in the 20 items, and the methodology used (Rasch modelling) is generally thought to be insensitive to this characteristic. An earlier study by Tuck (1995) employed several subscales such as Quality of Staff, School Climate and Academic Program, but no evidence was presented for the separate nature of these subscales. Salisbury et al (1997) demonstrated that parental satisfaction is associated with school effectiveness and student achievement. Parents' satisfaction is also found to be linked with other aspects of the school, especially the culture

and philosophy; the quality of the leadership and management; the behaviour and welfare of students; and the handling of issues such as bullying and harassment. For example, many features of a school's ethos improve once it has been identified as being an area of concern and appropriate action has been taken. However, in summary, the limited literature review of parent satisfaction with schooling provided little guidance as to the appropriate form such a survey should take.

Focus groups were conducted with teachers and parents in order to gain some insight into the dimensions of school satisfaction that were important to these groups. These aspects were prioritised and themes were identified which led to the development of eight possible areas (sub-domains) of assessment with a large pool of items (questions) to assess these sub-domains. The items were trial tested on a sample of 259 parents using an online survey. Parents were asked their level of agreement with each question on a five-point Likert scale. The survey was predominantly completed by mothers (66%), with a good spread across all possible years of schooling and types of schools (metropolitan, regional, primary, secondary, co-educational and single sex schools). However, the five-point scale proved to be unsatisfactory in that responses to items were concentrated at one rating point, 'agree'. This lack of discrimination at the item level meant that aggregate responses were difficult to distinguish, and precluded an assessment of relative performance between schools. The second problem stemming from the lack of variation within items was that co-variation between items was also restricted, making empirical validation of scale structure through correlations and factor analysis, difficult.

A follow-up study was conducted in 2007 with 8059 parents from 38 schools. This time, parents rated the questions on a scale of 1 to 10, from strongly disagree to strongly agree. The final Parent Satisfaction Survey consisted of 55 questions and focused on the eight key sub-domains: (1) Curriculum-Academic Program; (2) Quality of Teaching; (3) Learning Outcomes; (4) Pastoral Care; (5) Discipline and Safety; (6) Parent Involvement; (7) Resources; and (8) Year Transition. The number of parents per school varied substantially, although no school comprised 10% or more of the sample, and it was found that average satisfaction was not related to the size of school. Of the respondents, 72% were mothers, and parents of girls and boys were equally represented.

As for the pilot, parents tended to record high levels of satisfaction with all items. Average ratings were calculated in the eight sub-domains of satisfaction. These averages were more reliable according to the internal consistency coefficients. The relative standard deviation (the standard deviation as a percentage of the mean) ranged from a low of 19.3% to a high of 35.8% with a mean of 24.7%. This was a substantial improvement on the trial test data where the mean was 13.4%; and pointed to the value of increasing the rating categories from 5 to 10. Relationships between sub-domains (scales) were assessed by correlations, which were generally high. Items were allocated to sub-domains in order for scale scores to be determined. An exploratory factor model using maximum likelihood estimation of factors (with maximum likelihood estimation of missing data) and a promax (correlated factor) rotation was carried out using MPlus5. The exploratory five-factor model provided a reasonable fit, and so it was decided to continue with these sub-domains and question items with future Parent Satisfaction Surveys. However, the ten-point rating scale was later adjusted to an 11-point scale from 0 = no agreement to 10 = complete agreement to allow for a mid-point and easier reporting. This was because the average ratings, which were originally reported as percentage satisfaction to schools, had to be scaled by a linear

transformation to a mean of 75 and a standard deviation of 18 (the average of the raw standard deviations for each scale). This scaling meant that a number of scores became more than 100 and such scores were changed to the limiting values of 100.

2.2 Development of the Student Satisfaction Surveys

The Student and Year 12 Exit Student Satisfaction Surveys were developed in 2008 and followed a similar pattern to the Parent Satisfaction Survey. Many questions used in the Parent Survey were used and adapted for students. Additional questions, particularly pertaining to quality of teaching were sought from existing literature. There is limited research on measuring students' perceptions of their educational experience at school. However, some research has demonstrated that students who have a positive view of school are better motivated and achieve more according to their ability level than those who are not satisfied with school (Samdal, Nutbeam, Wold, & Kannas, 1998). Students hold valuable insight into the success of their school's operations and it is essential that their opinions be considered in developing, implementing, and sustaining a high quality education for current and future students (Calderon, Dobson, & Wentworth, 2000).

The Student Satisfaction Survey consists of three parts. The first part is the General Satisfaction Survey, which poses 50 questions evaluating the extent to which students in years 5 to 11 (aged between 10 and 17) believe effective school practices are apparent across the following nine sub-domains: (1) Academic Program; (2) Learning Outcomes; (3) Pastoral Care; (4) Personal Development/Leadership; (5) Discipline and Safety; (6) Resources; (7) School Ethos/Values; (8) Peer Relationships; and (9) Transition. The second part of the Student Satisfaction Survey evaluates students' perceptions of the quality of teaching at their school. These 38 questions evaluate the extent to which students believe effective school practices are apparent across five sub-domains: (1) Academic Rigour, (2) Feedback; (3) Teacher Knowledge; (4) Teacher Practice; and (5) Teacher-Student Rapport.

The third student survey, the Year 12 Exit Student Satisfaction Survey was also developed to more specifically look at the perceptions of students in their last year of school, since a few studies demonstrated that a supportive academic environment in year 12 is positively related to students' transition to life beyond secondary school (e.g., Khoo & Ainley, 2007; Pargetter et al., 1999; Tutton & Wigg, 1990). The transition experience is significantly influenced by students' perceptions of the quality of their secondary schooling. Khoo and Ainley (2007) found that student engagement and positive attitudes towards their school in their final years were strongly associated with participation and performance in further education and in the workforce.

2.3 Development of the Staff Satisfaction Survey

The Staff Satisfaction Survey was developed in 2011 using focus groups, questions from the existing *LEAD* Parent and Student Satisfaction Surveys and a pilot study of 894 teaching staff from four schools. Literature on school effectiveness and school improvement has consistently indicated that teachers are one of the most valuable resources available to a school. Staff satisfaction is also strongly associated with school effectiveness and student achievement. Some research confirms that the quality of what teachers know and do has the most impact on student learning (Hattie, 2003; Wenglinski, 2000;2002; Rowe, 2004). Hattie identified six major sources (variables) of explained variance in student learning outcomes

(namely students, teachers, schools, principals, home, and peer effects) and estimated their magnitude. He showed that in percentage terms, students' natural abilities account for about 50 per cent of the variance and this predicts achievement more than any other variable. The 'teacher effect' accounts for about 30 per cent of variance – the highest of the other variables. Hattie argues that the greatest influence on student learning is related to what teachers know and do, what they care about, and how they manage the learning for their students. The *L E A D* Staff Satisfaction Survey therefore aimed not only to look at staff perceptions of school effectiveness (such as academic programs, resources and pastoral care) but also to determine what workplace factors are important to ensure staff satisfaction.

Principal Axis Factoring (PAF) analysis using SPSS (Version 19) was used to analyse the 119 question items. Oblique rotation (Direct Oblimin with Kaiser rotation) was used to aid in the analysis of the factors because it allows factors to correlate, which is a condition more likely to occur in reality, particularly in opinion surveys. The Kaiser-Meyer-Olkin measure verified the sampling adequacy for the analysis (KMO = 0.981, which is 'superb' according to Field, 2009). Bartlett's test of sphericity, ($\chi^2=94472.92$, with $df=7021$ and $p<0.001$), indicated that correlations between items were sufficiently large for Principal Axis Factoring.

In order to test the assumptions about the various questions and where they would fit after the exploratory factor analysis, the 119 survey items were therefore divided into three subsets: 37 questions were identified as being 'workplace issues', 47 questions identified as being 'school-based issues' and 35 questions were being identified as 'teaching issues'. A Principal Axis factor analysis was conducted on each set. The resultant pattern matrices for each set were used to determine which survey items (or questions) fitted with which factor or sub-domain. The end result was a total of 16 factors: six factors (sub-domains) were determined for the workplace issues, seven factors (sub-domains) were determined for the school-based issues and three factors (sub-domains) for the teaching issues. Eight survey items were removed as having low factor loadings (below 0.2). The factors were tested for reliability and internal consistency using Cronbach's alpha (α). All of the factors had high reliabilities. It was decided for the purpose and ease of reporting, to discuss schools' results under two major themes, namely 'School-Based Domains' and 'Workplace Domains'. The three 'Teaching and Learning' domains were grouped under the 'School-Based Domains'.

The L E A D Project is still on-going and at the time of writing this paper, 174,366 parents, staff, and students from 159 schools have participated in 1,669 surveys. Whilst the majority of the schools involved have been independent (i.e., non-government) schools¹, they represent a highly diverse set of schools in terms of their size, religious affiliation, gender, educational philosophy, socio-economic status and locality. Most schools participate in the surveys biennially (with some undertaking them annually) and many have used the longitudinal data to effect changes and address issues that have emerged from the results.

¹ There are 9,393 schools in Australia. The schools are divided into three sectors, namely Government (70.9%), Catholic (18.3%) and Independent Schools (10.8%). https://www.is.vic.edu.au/independent/facts/pubs/fast_facts_2014.pdf and <http://isca.edu.au/about-independent-schools/>

3. Research Aims

Taking into consideration the basic characteristics of the school stakeholder surveys, particularly the breadth of the data gathered from parents, staff, students and members of the school boards, the availability of hard data (such as student results from national student tests and financial data) and the longitudinal nature of the measurements, this paper aims to: (a) explore the content and construct validity, the reliability and the internal consistency of the *LEAD* questionnaires and (b) provide suggestions to schools aiming to engage in a theory-driven and evidence-based approach towards improving their school effectiveness. The methodology and the main results of this study are presented below.

4. Research Methods

De-identified data were gathered from students, teaching staff, general (non-teaching) staff, parents and school board members taking the *LEAD* Surveys in 112 independent schools in Victoria, Australia, and followed a five-year longitudinal design from 2009 to 2013. The samples participating in this study consisted of 65,679 students in Years 5 to 12 (students aged from 10 to 18), 40,279 parents, 9,975 teaching staff and 3,816 general staff. The analyses were conducted in five distinctive domains in the following surveys included in the *LEAD* package: (a) General student satisfaction survey; (b) Student satisfaction with the quality of teaching; (c) Parent satisfaction survey; (d) Survey of teaching and general staff satisfaction with their schools and educational services (school-based domains); (e) Survey of teaching and general staff satisfaction with the schools as workplaces (workplace domains). The analyses evaluated the construct-validity, the reliability and the internal consistency of each of the aforementioned questionnaires, and the data were analysed through Confirmatory Factor Analyses approaches (CFA) using the AMOS and EQS software. For each survey, separate CFA analyses were conducted to identify the extent to which the theoretical models developed and used in the *LEAD* surveys came within acceptable fitting indices and parameters. For each model several fit indices have been estimated, such as the significance of the X^2 , the RMR, SRMR, GFI, CFI and RMSEA.

5. Data Analyses and Results

5.1 Content and Face Validity of the *LEAD* Surveys

Firstly, the *content validity* of the surveys was evaluated in collaboration with two Faculty members at the Faculty of Education, University of Cambridge, both of whom are considered to be leading experts in school improvement. This kind of validity refers to the extent to which the content of the measuring instruments is appropriate and relevant to the survey purpose. Content validity indicates whether the content reflects the complete range of the attributes under study and is usually undertaken by a number of experts (Pilot & Hunger 1999; DeVon et al. 2007; Antoniou, 2009). The conceptual framework of the surveys was found to be comprehensive and satisfactory in relation to contemporary research findings on school effectiveness research and especially with reference to multilevel models developed during recent years (e.g., Creemers & Kyriakides, 2008, Antoniou & Kyriakides, 2012); Creemers, 1994; Scheerens, 1992; Teddlie & Reynolds, 2000; Stringfield & Slavin, 1992). Some suggestions relating to the extent to which the questionnaire sub-domains could be

expanded are provided, such as the measurement of student and teacher behavior outside classrooms (during break-time) as part of the school ethos and culture; and the extent to which parents are given the opportunity to be involved in the teaching-learning process, when they have relevant expertise as part of a school collaboration policy.

Moreover, the *face validity* of the surveys was examined. Face validity indicates the extent to which the questionnaires appear to be appropriate to the survey purpose and content. It is the easiest validation process to undertake, but is the weakest form of validity (Haladyna 1999; Trochim 2001; DeVon et al. 2007). The two Faculty members mentioned previously, and two research assistants were asked to evaluate the face validity of the surveys. All provided positive comments on the appearance of the questionnaires in terms of feasibility, readability, consistency of style and formatting, and the clarity of the language used.

5.2 Construct Validity, the Reliability and the Internal Consistency of the *L E A D* Surveys

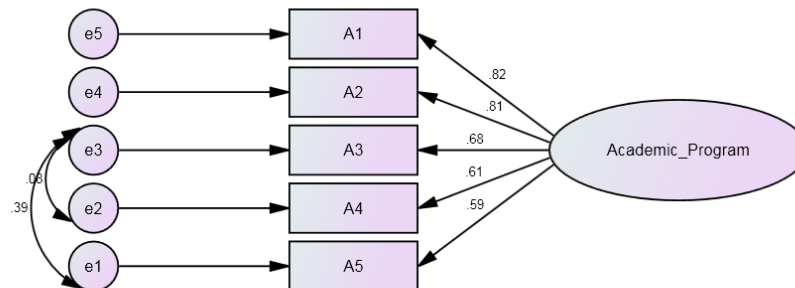
To evaluate the *construct-validity*, the reliability and the internal consistency of the *L E A D* Surveys, data were analysed through CFA Approaches using AMOS and EQS software. For each survey, separate CFA analyses were conducted to help identify the extent to which the theoretical models developed and used in the surveys came within acceptable fitting indices and parameters. To operationalise the various latent variables of the *L E A D* Surveys, the questionnaire items (predictors) were utilized. There are two major types of variables in structural equation modeling (SEM), namely, observed (indicator) variables and latent (construct) variables. Latent variables are not directly observable and hence they are inferred constructs, based on the observed variables that were selected to define each latent variable (Schumacker & Lomax, 2004: 196).

Cases with missing values posed an important challenge, because typical modelling procedures simply discard these cases from the analysis. When there are few missing values (very roughly, less than 5% of the total number of cases) and those values can be considered to be missing at random (that is, when a missing value does not depend upon other values), then the typical method of listwise deletion is relatively "safe". However, the percentage of the missing values (system missing) was much higher, because respondents were able to answer 'don't know' and 'not applicable' for each question. Thus, it was necessary to handle the missing values in the datasets using the Multiple Imputation procedure. The Multiple Imputation procedure provides analysis of patterns of missing data, geared toward eventual multiple imputation of missing values. That is, multiple versions of the dataset are produced, each containing its own set of imputed values. When statistical analyses are performed, the parameter estimates for all of the imputed datasets are pooled, providing estimates that are generally more accurate than they would be with only one imputation. This method of Multiple Imputation is generally considered to be superior to Single Imputation.

Having prepared the database for the analyses and implemented the Multiple Imputation procedure for missing values, CFA models were tested and compared to identify the final model with the optimum fit indices in each sub-domain of each survey. This procedure led to the development of fifty-four CFA models demonstrating the construct

validity of the questionnaire items in each sub-domain. An example is provided below, relating to the 'Academic Program' sub-domain of the General Student Satisfaction Survey.

Figure 2: CFA Results for Academic Program Sub-Domain of the *General Student Satisfaction Survey*



A CFA model designed to test the multidimensionality of the part of the General Student Satisfaction Survey measuring students' perceptions about the academic program of their school was used. The model hypothesised that the five variables (i.e. questionnaire items) could be explained by one factor and that each variable would have a non-zero loading on the factor it was designed to measure, and zero loadings on other factors. The findings of the first order factor SEM analysis generally affirmed the theory upon which the questionnaire was developed. Although the scaled chi-square for the one factor structure ($X^2=87.1$, $df=3$, $p<0.001$) was statistically significant, the values of RMSEA (0.031) and CFI (0.969) met the criteria for acceptable level of fit.

A similar approach was used for each sub-domain of the *LEAD* Surveys. As demonstrated in Table 1, nine models were generated for each construct included in the General Student Satisfaction Survey, five models were generated for each construct included in the Student Satisfaction with Quality of Teaching Survey and thirteen models were generated for each construct included in the Parent Satisfaction Survey. The analyses for the Staff Satisfaction Survey were split into three different sections: Ten CFA models were developed and tested for the Staff Satisfaction per School Domain, five models for the Staff Satisfaction per Workplace Domain and finally twelve models for the General (Non-teaching) Staff Satisfaction Survey.

In each of the models in Table1, to evaluate the construct validity and the reliability of the questionnaire sub-domains, several fit indices were estimated such as the significance of X^2 , the RMR, SRMR, GFI, CFI and RMSEA. In cases where the factor-loadings of questionnaire items were not found to be considerably high, alternative models which excluded the particular items were tested and compared with the original models. However, in all cases, it was found that the existing models yielded a better fit than alternative reduced models, thus, all questionnaire items were kept in the final CFA models. Such comparisons were made for several questionnaire items. Particularly, in the *Parent Satisfaction Survey*, comparisons were made for the Learning Outcome sub-domain, with and without item LO3 (i.e., 'My child is motivated to learn at this school' - with a factor loading of 0.49). The results of the analyses provided empirical support to the construct validity, the reliability and the

internal consistency of the questionnaires. However, the goodness of fit statistics was found to be somewhat marginal for the 'Quality of Teaching and Learning' sub-domain of the Teaching Staff Satisfaction Survey and for the 'Leadership and Morale' sub-domain of the General Staff Satisfaction Survey only.

Table 1: CFA models developed per questionnaire and sub-domain

Parents	Students		Teaching Staff		General Staff
Satisfaction Survey	General Satisfaction Survey	Satisfaction with Quality of Teaching Survey	Satisfaction Survey – per school domain	Satisfaction Survey – per workplace domain	Satisfaction Survey
Sub-Domain	Sub-Domain	Sub-Domain	Sub-Domain	Sub-Domain	Sub-Domain
Academic Program (6 items)	Academic Program (5 items)	Teacher Practice (9 items)	Resources and Offerings (5 items)	Feedback (5 items)	Resources and Offerings (4 items)
Quality of Teaching (7 items)	Learning Outcomes (5 items)	Teacher Knowledge (5 items)	Technology (5 items)	Goal Alignment (5 items)	Quality of Teaching and Learning (5 items)
Learning Outcomes (6 items)	Pastoral Care (4 items)	Feedback (7 items)	School Ethos/ Values (8 items)	Leadership and Morale (8 items)	Leadership and Morale (8 items)
Pastoral Care, Personal, Social Development (8 items)	Personal Development / Leadership (6 items)	Teacher Student Rapport (11 items)	Student Behaviour (8 items)	Staff Collaboration (9 items)	Student Behaviour (7 items)
Discipline and Safety (8 items)	Discipline and Safety (7 items)	Academic Rigour (6 items)	Discipline (4 items)	Professional Development (4 items)	Pastoral Care (8 items)
Parental Contact and Involvement (7 items)	Resources (4 items)		Pastoral Care (10 items)		School Ethos / Values (6 items)
Resources (5 items)	School Ethos and Values (5 items)		Parent Involvement (7 items)		Technology (4 items)
Transition (6 items)	Peer Relationships (5 items)		Learning Support (7 items)		Staff Collaboration (9 items)
	Transition (4 items)		Quality of Teaching and Learning (16 items)		Parent Involvement (4 items)
			Teaching Practice (9 items)		Feedback (5 items)
					Learning Support (4 items)
					Goal Alignment (4 items)

6. Conclusions and Suggestions for Further Studies

This paper is in line with the current approaches of utilising comprehensive SSE frameworks for identifying school needs and priorities for improvement. The comprehensive *LEAD* Suite of School Stakeholder Surveys and data analysis, providing benchmarked performance measures for school boards and senior leadership teams, has been presented and the construct-validity, the reliability and the internal consistency of each survey has been evaluated and supported through CFA approaches. The conceptual framework of the surveys was found to be comprehensive and satisfactory in relation to contemporary research findings on school effectiveness research, especially with reference to the multilevel models developed during the last years (e.g., Creemers & Kyriakides, 2008, Scheerens, 1992; Teddlie & Reynolds, 2000; Stringfield & Slavin, 1992). This is important because, although the importance of school self-evaluation has been reported extensively, the task of designing and developing instruments to collect data about each school factor is not an easy undertaking. To this end, the *LEAD* Surveys offer a validated package of tools, involving various groups of school stakeholders that can facilitate schools' efforts towards the identification of their needs and priorities for improvement.

This paper discusses the importance of schools evaluating their needs based on effectiveness and improvement theories and research, which have shown that effective improvement projects require school-level processes (Cuban 1998; Harris 2001). Teachers are considered an essential lever of change because change is explicit in their classrooms and their daily practices; but for effective school improvement, individual teacher initiatives are not enough. Teachers can succeed in achieving major changes in their classrooms with strong effects on student outcomes, but they cannot be expected to have a lasting impact on the school as an organisation. Improvement efforts initiated by one teacher will generally eventually disappear (e.g., when the teacher changes school) unless the school as an organisation sustains the efforts. At the same time, it is also recognised that innovations in classrooms need support at the school level and from all stakeholders for further incorporation and sustainability (Southworth 2002; Fresko et al. 1990).

This paper also stresses the need for promoting the design of school improvement projects that are based on a theoretical framework and measuring tools which have been systematically tested (Mosteller and Boruch 2002; Slavin 2002). The basic assumption supporting such initiatives is that improvement strategies must be based on evidence and therefore data should be collected to identify the improvement priorities of each school (Levin 2010). We also need to acknowledge that relying only on school stakeholders' experiences to develop their own evaluation tools can be limiting in terms of their school development and improvement (Britzman, 1991). This is because there appears to be little evidence from published literature that assisting school stakeholders to engage in any developmental or improvement program, without providing them with a validated framework to meet their personal needs and identify priorities for improvement, necessarily improves school effectiveness (Smith and Hatton 1992). Likewise, many studies stress the need for the provision of a well-researched and theory-based framework to form the basis of the improvement effort (e.g. Bierman et al. 2008; Buczynski and Hansen 2010; Domitrovich et al. 2008; Yoon et al. 2007).

Moreover, it is acknowledged that although schools may take steps towards identifying their needs and improvement priorities, this will not necessarily lead to the development and implementation of suitable educational policy aimed to improve educational outcomes. Many researchers (e.g. Creemers and Kyriakides 2006; Robertson and Sammons 1997; Teddlie and Reynolds 2000) have identified that an important constraint on the existing approaches of modelling and evaluating educational effectiveness is the fact that the whole process does not contribute significantly to the improvement of education. At this point it is also important to note that, despite a careful analysis of the failure to associate research and improvement effectively, some proposed school improvement strategies continue to attempt to combine the strong elements of research and improvement (Townsend 2007). Major elements of this combination are an emphasis on the evidence stemming from theory and research; a need to collect multiple data about the achievement of students and school processes on one hand; and an emphasis on the context of individual schools and, thereby, the development and implementation of programs by schools themselves, on the other (Creemers & Kyriakides, 2010).

Taking this into consideration, this paper aims to contribute to the development of the framework relating to the use of the *L E A D* Surveys, for school improvement purposes. From this perspective, drawing on the expertise of Independent Schools Victoria, analysis of the data and its results may help schools identify priorities for improving the functioning of specific factors. The schools involved in *The L E A D Project* are already engaged in this process. At the same time, although the Independent Schools Victoria team has an important role to play in analysing the data, the school stakeholders are also involved in this process and encouraged to identify the questions that need to be answered by the data analysis. In this way, the scope of the analysis is broadened and the special characteristics of the school are taken into account.

Suggestions for improving the *L E A D* Surveys could be drawn based on the above findings. For example, out of the 49 sub-domains analysed, the models of the two sub-domains with marginal fit indices could indicate the need to break down the one factor structure into two or more smaller components. Such components could subsequently be treated as the first-order factors formulating a second-order composite factor. This is also supported by research on teacher and school effectiveness in which, for example, quality of teaching comprises of many sub-factors (first-order factors). Particularly, in a number of studies measuring school effectiveness, quality of teaching consists of eight factors (see Creemers & Kyriakides, 2006;2008, Kyriakides, Creemers & Antoniou, 2009; Antoniou and Kyriakides, 2011;2013a;2013b). Moreover, suggestions could be drawn relating to the sampling methodology used in the *L E A D* Surveys to change the data collection procedures and facilitate multilevel modelling analyses of the data. For example, as is the case in of the *L E A D* Surveys, when the measurements relate to the same variable but measured at different points in time, multilevel analysis can be employed to model individual school growth patterns. The main advantage of multilevel analysis for this purpose is its flexibility and capability to deal with unbalanced data. Multilevel analysis does not require the same number of measurements for each case and can easily handle data with incomplete records on the outcome measures.

Further studies could expand the research framework to identify links between the latent structures of the surveys and other factors such as school improvement over time; school and student socio-economic status; student achievement; as well as other variables.

Taking into consideration the basic characteristics of the surveys, particularly the breadth of data gathered from parents, staff, students and members of the school boards; the availability of objective (such as student tests and financial data) and subjective data (i.e., stakeholders' perceptions); and the longitudinal nature of the measurements, the analyses could have important implications for understanding the relationship between school stakeholders' perceptions and school effectiveness and the day-to-day planning and policy development at the school or system level.

Acknowledgements

The authors of this paper wish to acknowledge the following people:

Michelle Green, Chief Executive of Independent Schools Victoria, for her vision in the development and continuation of *The L E A D Project*.

Assoc. Prof Richard C. Bell and Sophie Keele from The University of Melbourne School of Behavioural Science for their initial contributions to the parent and student surveys.

Matthew Ryan, a Canberra economist, for the conceptual and architectural development of *The L E A D Report*.

Current and past research staff members at Independent Schools Victoria.

The Principals, parents, students and staff at the Member schools of Independent Schools Victoria, for their continued assistance and support for *The L E A D Project*.

Information about Independent Schools Victoria, Australia can be accessed here:
www.is.vic.edu.au.

Information about the *L E A D* suite of surveys can be accessed at:
<http://www.is.vic.edu.au/services/supporting-schools/lead.htm>

References:

- Antoniou, P. & Kyriakides, L. (2013a). A Dynamic Integrated Approach to Teacher Professional Development: Impact and Sustainability of the Effects on Improving Teacher Behavior and Student Outcomes. *Teaching and Teacher Education*, 29 (1), 1-12.
- Antoniou, P., Kyriakides, L., & Creemers, B.P.M. (2013b). The Impact of School Policy and Stakeholders' Actions on Student Learning Outcomes: A Longitudinal Study. *Paper presented at the 2013 British Educational Leadership, Management and Administration Society (BELMAS) annual meeting, Edinburgh, Scotland.*
- Antoniou, P., Kyriakides, L., & Creemers, B.P.M. (2012). Improving Teacher Effectiveness. In Creemers, B.P.M., & Kyriakides, L. (Eds). *Improving Quality in Education: Dynamic Approaches to School Improvement*. London and New York: Routledge.
- Antoniou, P. & Kyriakides, L. (2011). The Impact of a Dynamic Approach to Professional Development on Teacher Instruction and Student Learning: Results from an Experimental Study. *School Effectiveness and School Improvement*, 22 (3), 291 – 311.
- Antoniou, P. (2009). A group-randomization study investigating the effectiveness of different approaches to teacher professional development: Preliminary results. *Paper presented at the annual meeting of the International Congress on School Effectiveness and School Improvement (ICSEI), Vancouver: Canada.*
- Barber, M. (1996). *The learning game: arguments for an education revolution*. London: Victor Gollanz.
- Bierman, K. L., R. L. Nix, M. T. Greenberg, C. Blair, & C. E. Domitrovich. (2008). Executive Functions and School Readiness Intervention: Impact, Moderation, and Mediation in the Head Start REDI Program. *Development and Psychopathology*, 20, (3).
- Britzman, D.P. (1991). *Practice Makes Practice: A Critical Study of Learning to Teach*. Albany: State University Press.
- Broadhead, P., Cuckle, P., Hodgson, J., & Dunford, J. (1996). Improving Primary Schools through School Development Planning. Building a Vision, Exploring the Reality. *Educational Management Administration & Leadership*, 24(3), 277-290.
- Buczynski, S., & C. B. Hansen. (2010). Impact of Professional Development on Teacher Practice: Uncovering Connections. *Teaching and Teacher Education*, 26 (3): 599-607.
- Calderon, A. J., Dobson, I. R., & Wentworth, N. (2000). Recipe for success: Year 12 subject choice and the transition from school to university. *Journal of Institutional Research*, 9(1), 111-123.
- Caldwell, B.J. (2003). A theory of learning in the self managing school. In A. Volansky & I. Friedman (Eds.), *School-based management: An international perspective* (pp. 93–116). Jerusalem: State of Israel, Ministry of Education.
- Canner, J. (1985). *School improvement project: 1983-1984. Fifth annual process assessment*. New York: New York City Board of Education.

- Constable, H. (1994). Three arenas of tension: Teachers' experience of participation in school development planning. In D.H. Hargreaves & D. Hopkins (Eds.), *Development planning for school improvement* (pp. 91-99). New York: Cassell.
- Creemers, B.P.M., Kyriakides, L. & Antoniou, P. (2013). A Dynamic Approach to School Improvement: Main Features and Impact. *School Leadership and Management*, 33 (2), 114-132.
- Creemers, B. P. M. & Kyriakides, L. (2012). Improving quality in education: dynamic approaches to school improvement. London: Routledge.
- Creemers, B. P. M., & Kyriakides, L. (2010). Using the Dynamic Model to Develop an Evidence-Based and Theory-Driven Approach to School Improvement. *Irish Educational Studies* 29, (1): 5-23.
- Creemers, B.P.M., and L. Kyriakides (2008). The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools. London: Routledge.
- Creemers, B. P. M. & Kyriakides, L. (2006). Critical analysis of the current approaches to modelling educational effectiveness: the importance of establishing a dynamic model. *School Effectiveness and School Improvement*, 17(3), 347–366.
- Creemers, B.P.M. (1994). *The effective classroom*. London: Cassell.
- Cuban, L. (1998). How Schools Change Reforms: Redefining Reform Success and Failure. *Teachers College Record*, 99 (3): 453-477, <http://www.tcrecord.org/content.asp?contentid=10273>.
- Desimone, L.M., Smith, T.M., Hayes, S.A., & Frisvold, D. (2005). Beyond Accountability and Average Mathematics Scores: Relating State Education Policy Attributes to Cognitive Achievement Domains. *Educational Measurement: Issues and Practice*, 24(4), 5–18.
- DeVon, H. A., Block, M. E., Moyle-Wright, P., Ernst, D. M., Hayden, S. J., Lazzara, D. J. et al. (2007). A psychometric Toolbox for testing Validity and Reliability. *Journal of Nursing scholarship*, 39 (2), 155-164.
- Devos, G. (1998). Conditions and caveats for self-evaluation: the case of secondary schools, *Paper presented at the Annual Meeting of the American Educational Research Association*, San Diego: (ERIC Document Reproduction Service No. ED421493).
- Domitrovich, C. E., S. D. Gest, S. Gill, K. L. Bierman, J. Welsh, & D. Jones. (2008). Fostering High Quality Teaching in Head Start Classrooms: Experimental Evaluation of an Integrated Curriculum. *American Educational Research Journal*, 46 (2): 567-597.
- Field, A (2009) *Discovering Statistics Using SPSS (and sex and drugs and rock 'n roll)*, 3rd Edition, Sage Publications Ltd., London, California, New Delhi, Singapore
- Flessa, J.J. (2012). Principals as middle managers: School leadership during the implementation of primary class size reduction policy in Ontario. *Leadership and Policy in Schools*, 11(3), 325-343.

- Freiberg, H. J. (Ed.) (1999). *School climate: measuring, improving and sustaining healthy learning environments*. London: Falmer Press.
- Fresko, B., N. Robinson, A. Friedlander, J. Albert, and N. Argaman. (1990). Improving Mathematics Instruction and Learning in the Junior High School: An Israeli Example. *School Effectiveness and School Improvement*, 1 (3): 170-188.
- Fullan, M. (2001). *The new meaning of educational change* (3rd ed). New York: Teachers College Press.
- Goldstein, H. (1995). *Multilevel statistical models* (2nd ed). London: Edward Arnold.
- Haladyna, T. (1999). *Developing and Validating multiple-choice test items*. New Jersey: Lawrence Erlbaum.
- Harris, A. (2001). Building the Capacity for School Improvement. *School Leadership & Management*, 21 (3): 261_270. doi:10.1080/13632430120074419.
- Hattie, J.A. (2003). *Teachers make a difference: What is the research evidence?* Background paper to invited address presented at the 2003 ACER Research Conference, Carlton Crest Hotel, Melbourne, Australia, October 19-21, 2003. Cited in Rowe, KR (2004), see below.
- Hofman, R.H., Hofman, W.H., & Gray, J.M. (2010). Institutional contexts and international performances in schooling: Comparing patterns and trends over time in international surveys. *European Journal of Education*, 45(1), 153-173.
- Hoy, W. K. & Miskel, C.G. (2001). *Educational administration: theory, research and practice*. Singapore: McGraw-Hill.
- Khoo, S. T., & Ainley, J. (2007). The Influence of School Attitudes, Achievement, and Educational Intentions on School Completion and University Participation, ACER (Australian Council for Educational Research).
- King, J.A., & Bond, T.G. (2003) measuring client satisfaction with public education I: Meeting competing demands in establishing statewide benchmarks. *Journal of Applied Measurement*, 4(2), 111-123.
- Kyriakides, L. & Campbell, R. J. (2004). School self-evaluation and school improvement: a critique of values and procedures. *Studies in Educational Evaluation*, 30(1), 23–36.
- Kyriakides, L., Creemers, B.P.M. & Antoniou, P. (2009). Teacher behaviour and student outcomes: Suggestions for research on teacher training and professional development. *Teaching and Teacher Education*, 25 (1), 12-23.
- Levin, B. (2010). Governments and education reform: Some lessons from the last 50 years. *Journal of Education Policy*, 25(6), 739–747.
- MacBeath, J. (1999). *Schools must speak for themselves: The case for school self-evaluation*. London: Routledge.

- Mosteller, F., & R. Boruch. (2002). *Evidence Matters: Randomized Trials in Education Research*. Washington, DC: Brookings
- Muijs, D. & Reynolds, D. (2001). *Effective teaching: evidence and practice*. London: Sage.
- Nir, E.A., & Ben Ami, T. (2005). School-Parents Relationship in the Era of School-Based Management: Harmony or Conflict? *Leadership and Policy in Schools*, 4(1), 55-72.
- Pargetter, R., McInnis, C., James, R., Evans, M., Peel, M., Dobson, I. (1999). Transition from Secondary To Tertiary: A Performance Study. *Department of Education Training and Youth Affairs (DETYA) Higher Education Series*, No. 36.
- Pilot, D., & Hunger, B. (1999). *Nursing research: principals and methods*. Philadelphia: Lippincott Williams & Wilkins.
- Robertson, P.J., Wohlstetter, P., & Mohrman, S.A. (1995). Generating curriculum and instructional innovations through school-based management. *Educational Administration Quarterly*, 31(3), 375-404.
- Robertson, P., & P. Sammons. (1997). Improving School Effectiveness Project: Understanding Change in Schools. *Paper presented at the British Educational Research Association Annual Conference*, September 11-14, University of York, UK.
- Rowe, K.J. (2004). The importance of teaching: Ensuring better schooling by building teacher capacities that maximize the quality of teaching and learning provision – implications of findings from the international and Australian evidence-based research. *Background paper to invited address presented at the Making Schools Better Summit Conference*, Melbourne Business School, the University of Melbourne, 26-27 August 2004. Available at: www.acer.edu.au/research/programs/learningprocess.html
- Rosenkvist, M.A. (2010). Using Student Test Results for Accountability and Improvement: A Literature Review. *OECD Education Working Papers*, No. 54. Paris: OECD Publishing.
- Salisbury, D. F., Branson, R. K., Altreche, W. I., Funk, F. F., & Broetzmann, S. M. (1997). Applying customer dissatisfaction measures to schools: You better know what's wrong before you try to fix it. *Educational Policy*, 11 (3), 286-308.
- Samdal, O., Nutbeam, D., Wold, B., & Kannas, L. (1998). Achieving health and educational goals through schools – a study of the importance of the school climate and the students' satisfaction with school. *Health Education Research*, 13(3), 383-397.
- Scheerens, J. & Bosker, R. J. (1997). *The foundations of educational effectiveness*. Oxford: Pergamon.
- Scheerens, J. (1992). *Effective schooling: Research, theory and practice*. London: Cassell.
- Schumacker, R.E., & Lomax, R.G. (2004). *A Beginner's Guide to Structural Equation Modelling* (2nd ed). Lawrence Erlbaum Associates Inc., Mahwah, NJ.
- Slavin, R. E. (2002). Evidence-Based Education Policies: Transforming Educational Practice and Research. *Educational Researcher*, 31 (7): 15-21.

- Smith, M.S., & O'Day, J. (1991). Systemic school reform. In S. Fuhrman & B. Malen (Eds.), *The politics of curriculum and testing* (pp. 233-268). Philadelphia: Falmer Press.
- Smith, L., & N. Hatton. (1992). Towards reflection in teacher education. What counts as Evidence? Paper Presented at the Annual Conference of the Australian Association for Research in Education, November, in Deakin University, Australia.
- Southworth, G. (2002). Instructional Leadership in Schools: Reflections and Empirical Evidence. *School Leadership & Management*, 22 (1): 73-91.
- Spillane, J.P. (2005). Primary school leadership practice: How the subject matters. *School Leadership and Management*, 25(4), 383–397.
- Stringfield, S.C., & Slavin, R.E. (1992). A hierarchical longitudinal model for elementary school effects. In B.P.M. Creemers & G.J. Reezigt (Eds.), *Evaluation of Educational Effectiveness* (pp. 35-69). Groningen: ICO.
- Teddlie, C., & Reynolds, D. (2000). *The International Handbook of School Effectiveness Research*. London: Falmer Press.
- Townsend, T. (2007). *International Handbook of School Effectiveness and Improvement*. Dordrecht: Springer.
- Trochim, W.M.K. (2001). *The Research Methods Knowledge Base*. Cincinnati: Atomic Dog.
- Tuck, K.D. (1995) Parent Satisfaction and Information (A Customer Satisfaction Survey). District of Columbia Public Schools, Washington, DC. Research Branch. ERIC Document ED401326
- Tutton, P. J. M. & Wigg, S. J. (1990). The Influence of Last Secondary School Attended, Subjects Taken in Last Year of Secondary Education and Gender on Preclinical Performance of Medical Students, *Australian Journal of Education*, 34(2), 168-173.
- Wenglinski, H. (2000). How Schools Matter: The Link Between Teacher Classroom Practices and Student Academic Performance, accessed on July 12 2005: <http://epaa.asu.edu/epaa/v10n12/>
- Wenglinski, H (2002). Teacher classroom practices and student performance: How schools can make a difference, accessed on 12 July 2005: www.ncrel.org/gap/library/text/26teachersmake.htm
- Whitty, G., & Edwards, T. (1998). School choice policies in England and the United States: an exploration of their origins and significance. *Comparative Education*, 34(2), 211–227.
- Vitaska, S. (2008). *Strong leaders, strong schools: 2007 state laws*. Denver, CO: National Conference of State Legislatures.
- Yoon, K. S., M. Garet, B. Birman, & R. Jacobson. (2007). Examining the Effects of Mathematics and Science Professional Development on Teachers' Instructional Practice: Using Professional Development Activity Log. Washington, DC: Council of Chief State School Officers.