Paper Code: 00419

Measurement and Modeling: Sequential Use of Analytical Techniques in Education Contexts

Kelsey Halbert, James Cook University, kelsey.halbert@jcu.edu.au

ABSTRACT

This paper presents a case study of methodology that combines sequential use of psychometric and traditional measurement techniques. Such combined techniques have been used in educational contexts and the literature indicates that this may provide a more rigorous approach than is possible using only one measurement approach (Cavanagh & Romanosky in Waugh, 2005).

The educational context being explored is reasoned risk-taking in decision-making by school principals. In this paper, risk-taking is defined as when decisions are made that are not compliant with the regulatory framework, the primary governance mechanism for public schools in Western Australia. Such decisions involve risk as principals may be exposed to criticism for non-compliance with established policy when negative outcomes arise from decision-making. This creates a dilemma for principals who need to be able to respond to the locally identified needs within a school, and simultaneously comply with all State and Commonwealth departmental requirements.

Data was collected through the survey of a stratified random sample of principals in 253 Western Australian government schools. The questionnaire included measures of both attitude and behaviour of principals to determine whether reasoned risk-taking by school principals is a consequence of their perceptions of the governance mechanism of the regulatory framework, the experience of individual principals and the characteristics of key stakeholders within the school community.

Rasch measurement has been used to analyse the data and create a robust measurement scale. This psychometric technique provides current world best practice in the creation of linear scales in the human sciences. A theoretical model of factors impacting on reasoned risk-taking in decision-making was developed and tested used Partial Least Squares (PLS) structural equation modeling. This analysis provides evidence of the effect of the concepts included in the model and consequently insight into governance structures, characteristics of schools and principals that influence decision-making in schools.

BACKGROUND

There is currently a conflict between the rhetoric of decentralisation and external requirements in Australian schools. Since the 1960's the political climate of Western nations and demands of cultural minorities and women for increased participation have contributed to the rise of school-based decision-making and management as an administrative strategy in education (Seddon, Angus and Poole, 1990, pp.29-41). In a comparison of 19 countries, the 2004 OECD report found that in 14 countries, decisions were being made at a more decentralised level in 2003 than in 1998. However, Australia was found to be one of the countries with the most centralised educational decision-making (Caldwell, 2006, p.65). This is despite research and government reports, such the Karmel Report (1973, p.10), recommending that Australian schools move towards a more decentralised form of management. A commitment to decentralisation and devolution of authority in education was made at a national level following the election of the Australian Labor Party in 1983 (Caldwell, 1990, p.5) and national and state government initiatives over recent years are still tending to move in this direction (Eacott, 2009; Department of Education and Training, 2009).

The impact of management demands and the requirements of central education authorities in constraining innovation in schools has been an issue of debate for many years (Sarason, 1982; Bennis and Nanus, 1985; Sergiovanni, 2000; Starr, 2008). Principals have the dual task of being instructional leaders to ensure that students attain achievement standards and simultaneously lead and manage the organisation of the school. As each school is a component of a larger organisation there are requirements imposed from the organisational executive regarding both educational and business aspects of the leadership role of principals. Sergiovanni (2000, p.166) observes that school professionals don't have a "high tolerance for bureaucratic rituals" as they are often responding to a range of competing stakeholder demands in a politically exposed environment. Fullan (1993, p.22) concludes that "you can't mandate what matters" as educational goals are complex and require discretionary judgement. This view is supported by the research on school effectiveness and school improvement that has been consistently supportive of school-based decision-making and management (Caldwell, 1990, p.19).

Reasoned Risk-Taking in Decision-Making

The relationship of reasoned risk-taking to governance has been studied in a range of business contexts. Carpenter, Pollock and Leary (2003), Wiseman and Gomez-Meija (1998) and Wiseman et al (2000) have tested models of reasoned risk-taking and its relationship to governance mechanisms. Within these studies agency and behavioural perspectives have been used to develop theories of risk-taking where the nature of risk-taking is a consequence of governance mechanisms and stakeholder characteristics including the experience of management.

Much of the literature looking at models of strategic decision-making and risk-taking behaviour relates to studies that have been conducted within the context of business environments. For example, Carpenter, Pollock and Leary (2003) provided some evidence of the critical role of boards and management, as a component of corporate governance, on strategies involving reasoned risk-taking and that the risk attitudes and experience of individual principals and agents can influence reasoned risk-taking strategies of firms. This paper seeks to apply these theories to the public sector environment, specifically the context of reasoned risk-taking by principals in government schools. Decision-making in the public sector is controlled by the governance mechanism of regulatory frameworks that include legislation, regulations and departmental polices. Reasoned risk-taking occurs when decisions are made that are not compliant with the regulatory framework. This paper proposes the use of agency (Hoskisson et al. 1999, p. 435; Rumelt, Schendel & Teece, 1991, p.15) and behavioural (Wiseman & Gomez-Mejia, 1998) perspectives to determine whether the reasoned risk-taking by school principals is a consequence of their perceptions of the governance mechanism of the regulatory framework, the experience of principals and the characteristics of key stakeholders within the school community.

METHODS

This paper develops and validates a hypothesised model of risk-taking in decision-making by school principals. The model, and the associated questionnaire to measure constructs composing it, were developed with reference to the literature and conclusions from a qualitative familiarisation study. The familiarisation study involved interviews with government school principals and district office administrators across the state of Western Australia.

Data Collection

Preliminary data was collected through face-to-face interviews with principals and district directors in each of the 16 district education offices across Western Australia. In addition, interviews were conducted with key stakeholder groups. This data was used, in conjunction with the literature review, to develop the research model below.

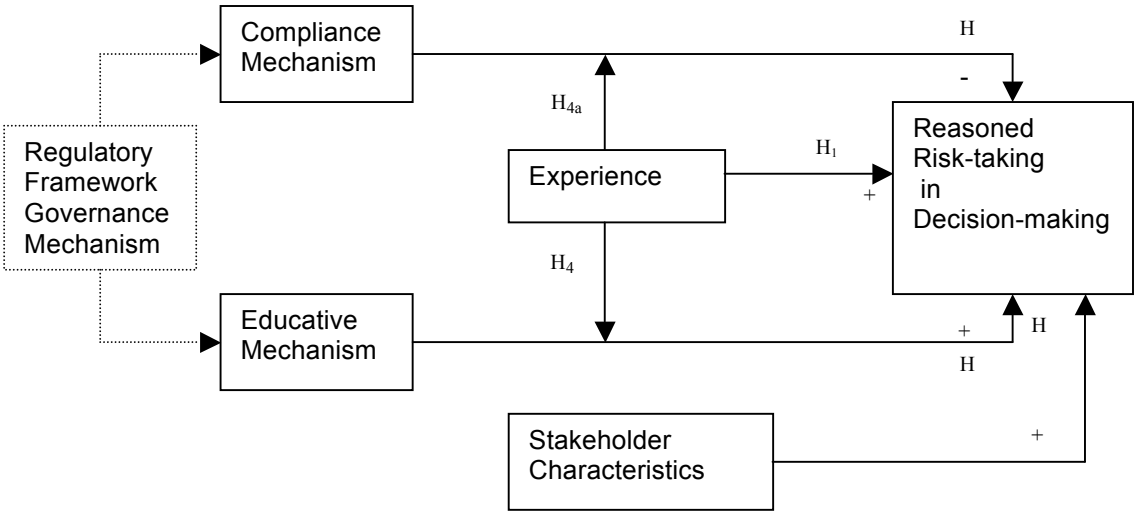


Figure 1: Research Model

Measurement items were developed for each of the constructs included in the research model. Measurements in existing studies related to business environments were not transferable to an educational context. A survey questionnaire was then developed to measure the constructs included in the research model. The survey questionnaire was piloted with a small sample of principals from both metropolitan and remote primary and secondary schools. The refined version of the survey questionnaire was then used to collect confirmatory data to test the research model.

ANALYSIS

A combination of Rasch and traditional statistical techniques, including structural equation modeling, were applied to test the measurement properties of the questionnaire and the hypothesised model. The choice of analytic techniques in this paper was made to provide the strongest evidentiary base to support the validity of the theoretical model. Cavanagh and Romanoski (2005, p.67) point out that "although these techniques are based upon different psychometric traditions, each has utility in specific aspects of hypothesis testing and measurement".

The data analysis undertaken involved four procedures. Firstly, a preliminary statistical analysis of each of the items in the questionnaire including the demographic variables of interest was conducted using SPSS version 12 (2003). Following this, a Rasch analysis was conducted with RUMM version 2020 (2005) to explore the psychometric properties of the measurement instrument

used to collect the data. Having established that the questionnaire provided a valid and reliable scale of measurement, an analysis of the model incorporating these constructs was conducted using exploratory factor analysis. Finally, the hypotheses were tested using structural equation modeling. The structural equation modeling was conducted using Partial Least Squares (PLS) version 3.0. (Chin, 2001).

Structural equation modeling and other path analysis techniques have been more frequently used in educational based research in recent years. Fidler (2001, p.54) comments that there are opportunities for greater use of such techniques in understanding schools and their effectiveness. This paper combines structural equation modeling with Rasch analysis. Cavanagh and Romanoski (2005 p.68) contend that the Rasch methodology allows identification of measurement errors due to item or person misfit during scale construction, allowing the errors to be minimised prior to testing of the model with traditional techniques. Structural equation modeling includes an assessment of the measurement model as a component of the analysis. However, the presence of errors of measurement and their influence on the fit of the data to the model are only revealed after the fit statistics have been estimated. Cavanagh and Romanoski (2005) posit that the presence of such errors could prevent model confirmation. Following conduct of the Rasch analysis, items identified as demonstrating errors were removed from the analysis before undertaking the structural equation modeling analysis. This paper details the Rasch analysis undertaken before the model was tested with structural equation modeling.

Rasch Analysis

Rasch analysis (Rasch, 1960/1980; Andrich, 1988 & 1989; Lunz & Linacre, 1998) was the methodology used to examine the psychometric properties of the questionnaire data collected to test the hypothesised model of reasoned risk-taking in decision-making by school principals. The procedure involves scaling the results of principals on each item in the questionnaire relative to their responses on the other items. The procedure for analysing differential performance uses the principles of latent trait theory. The model requires that there is a single latent trait which governs the responses of all persons to all items. In this paper this trait would be reasoned risk-taking in decision-making. This component of the analysis aimed to produce a measurement scale of the attitudes and behaviours of school principals towards risk-taking in decision-making.

The use of Rasch analysis has several strengths in relation to traditional statistical techniques. Application of the Rasch measurement model requires that variables be measured in common units and also that persons and items are positioned on the same interval scale. In addition, the calibration process for Rasch analysis ensures that the scales developed are linear. “Testing the psychometric properties of data from a scale using classical techniques will not reveal errors due to item disordering” (Cavanagh & Romanosky in Waugh, 2005, p.68). Cavanagh and Romanosky (Waugh, 2005, p.68) and Cavanagh and Waugh (2004) indicate that measurement errors due to disordered Likert scale responses are undetected by traditional techniques, whereas items where respondents have not answered consistently or logically can be detected using Rasch and omitted from further analyses. Use of the measurement stage of structured equation modeling analyses provides tests of the reliability of the items in the model and information about their functioning. However, Cavanagh and Romanoski (Waugh, 2005, p.68) point out that SEM will not reveal errors due to item disordering or test whether the respondents have answered survey items consistently and logically. Rasch phychometric scale analysis provides insight into these additional sources of measurement error.

The questionnaire was analysed by using RUMM 2020 (Andrich, Sheridan & Luo, 2005) to validate the data on the items in the questionnaire and persons responding to it and to develop an interval scale comprised of items that were determined to fit the model. The algorithm in the program (Andrich, Sheridan & Lyne, 1991) uses a pairwise procedure, and providing there is no overlap in the response patterns, it handles missing data routinely. Therefore it was able to handle the items to which a principal did not respond.

Results of the Rasch Analysis

In this analysis centralised item thresholds were calculated. The item thresholds are shown in Table 1. These thresholds were scrutinised for items with disordered thresholds. Identified items were discarded from further analysis as the existence of disordered thresholds indicates that the items were not operating logically or consistently in regard to responses provided on the Likert scale. Disordered thresholds were found for items 8 Location; 9 School Type; 11 and 12 the proportions of Indigenous and ESL students. These items were retained as they were open response quantitative items and were not based on a Likert scale. However, questions 13, 25, 29, 33, 38, 43, 56, 57, 59, 60, and 61 were discarded on the basis of disordered thresholds as they were not operating as expected. Data from the items with ordered thresholds were retained for further analysis.

Table 1: Centralised Thresholds

Question	Location	CenThr 1	CenThr 2	CenThr 3	CenThr 4	CenThr 5	CenThr 6
Location	-0.13984	-0.77	1.441	-0.671			
School Type	1.00623	0.376	-1.158	1.456	-0.673		
Indigenous	0.251236	0.774	0.273	-0.262	-0.651	-0.715	-0.273
ESL	0.220794	0.727	-0.066	-0.345	-0.294	-0.096	0.066
13	-3.90127	0.207	-2.828	2.621			
14	0.282702	-2.089	0.465	1.624			
15	-0.57375	-1.309	-0.752	2.061			
16	0.268026	-2.48	0.16	2.32			
17	-0.27838	-2.152	-0.54	2.692			
18	-0.72241	-1.672	-0.546	2.218			
19	-1.02836	-1.753	0.226	1.527			
20	-1.96758	-3.933	1.151	2.781			
21	-2.00502	-2.507	-0.199	2.706			
22	-0.5849	-3.513	0.891	2.622			
23	-2.00582	-3.793	1.194	2.599			
24	0.02609	-1.231	-0.138	1.368			
25	-0.83349	-0.8	-1.511	2.311			
26	-1.73115	-4.342	0.529	3.813			
27	-1.01804	-1.454	-0.215	1.669			
28	-0.31088	-1.365	-0.392	1.757			
29	1.075536	-0.612	0.505	0.107			
30	-0.16479	-2.309	0.45	1.858			
31	-0.62698	-1.893	-0.103	1.997			
32	-0.1273	-1.904	-0.876	2.779			
33	-0.63968	-0.442	-1.412	1.853			
34	-1.96805	-4.851	2.318	2.533			
35	0.083649	-1.604	-0.994	2.599			
36	1.101064	-2.157	-0.579	2.737			
37	-0.62627	-2.37	-0.342	2.712			
38	-0.02066	-1.992	1.251	0.741			
39	0.207149	-1.964	0.177	1.788			
40	-0.22025	-1.531	-0.583	2.115			
41	0.471836	-2.147	-0.336	2.483			
42	1.17015	-2.467	0.785	1.682			
43	-0.32486	-0.728	-1.008	1.736			
44	-0.46365	-1.698	-0.497	2.195			
45	-0.3569	-3.491	1.15	2.341			

46	-0.35659	-3.702	0.341	3.361
47	-0.85565	-2.782	0.622	2.159
48	-0.68758	-2.343	0.642	1.701
49	0.547331	-2.042	0.012	2.029
50	-0.71556	-1.462	-0.128	1.59
51	-0.6908	-3.019	0.718	2.301
52	-0.68558	-2.814	0.525	2.289
53	-1.92285	-3.682	0.231	3.451
54	1.33411	-2.845	0.72	2.125
55	0.53637	-0.921	0.032	0.888
56	4.523194	-3.154	2.9	0.254
57	1.947693	0.355	4.709	-5.064
58	0.239727	-1.979	0.123	1.855
59	0.853383	-0.999	0.528	0.471
60	1.787935	-1.05	0.96	0.09
61	1.143501	-1.356	0.77	0.587
62	1.78908	-2.142	0.586	1.556
63	1.788379	-1.863	0.354	1.509
64	2.991461	-2.433	0.235	2.197
65	2.908249	-2.511	-0.12	2.631

Closer scrutiny of the category response frequencies for items 13, 25, 33, 56, 57 and 60 showed that these items were poorly targeted for this group of respondents. All respondents either agreed or strongly agreed to items 13 and 56. Very few respondents disagreed with items 25, 33 or 57 and few agreed with item 60. As a consequence these items failed to adequately discriminate between respondents.

The retained items were subsequently examined for low residuals and high Chi Square probability. The residual for an item is the difference between the actual response and the expected response to the item as predicted by the measurement model. A low residual of $<\pm 2.0$ indicates that the item fits the model, whereas a high residual shows poor fit to the model. The Chi Square estimates the probability that the item's data fit the model whereas a probability value with $p<0.05$ shows poor fit to the model. Item 12 had a high residual and a high Chi Square and was discarded from further analysis. The results of this analysis are presented in Table 2. The individual item-fit statistics also show that the majority of items fit the model.

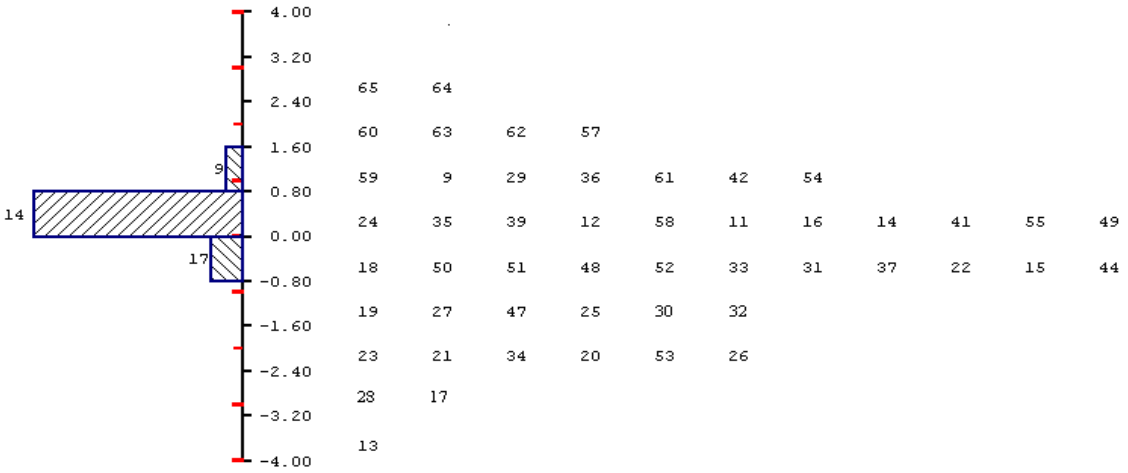
Table 2: Individual Item Fit

Item	Location	SE	FitResid	DF	ChiSq	DF	Prob
8	-0.14	0.081	1.081	133.36	2.925	2	0.231623
9	1.006	0.079	0.483	133.36	0.256	2	0.879644
11	0.251	0.042	1.57	130.48	12.173	2	0.002274
12	0.221	0.043	2.786	118.01	29.971	2	0
13	-3.901	0.225	0.233	134.32	0.369	2	0.831437
14	0.283	0.12	-0.193	131.44	9.962	2	0.006867
15	-0.574	0.137	0.294	131.44	0.216	2	0.897691
16	0.268	0.133	0.226	131.44	0.861	2	0.650121
17	-0.278	0.148	0.361	130.48	0.469	2	0.790941
18	-0.722	0.142	0.374	128.56	0.649	2	0.722771
19	-1.028	0.119	0.272	134.32	1.291	2	0.524335
20	-1.968	0.126	0.512	133.36	2.182	2	0.33594
21	-2.005	0.151	0.099	133.36	0.396	2	0.820399
22	-0.585	0.133	-0.304	133.36	2.176	2	0.33689
23	-2.006	0.125	0.375	131.44	1.224	2	0.542164
24	0.026	0.107	0.841	128.56	1.802	2	0.406246

25	-0.833	0.159	0.143	133.36	2.202	2	0.332605
26	-1.731	0.163	0.435	128.56	1.128	2	0.568813
27	-1.018	0.128	-0.17	131.44	2.63	2	0.268436
28	-0.311	0.122	0.32	127.6	0.931	2	0.627787
29	1.076	0.103	0.728	131.44	12.744	2	0.00171
30	-0.165	0.124	-0.149	125.68	4.312	2	0.115817
31	-0.627	0.128	-0.147	128.56	4.677	2	0.096451
32	-0.127	0.155	0.247	126.64	4.634	2	0.098558
33	-0.64	0.142	0.095	129.52	0.06	2	0.970619
34	-1.968	0.109	-0.757	132.4	5.459	2	0.065263
35	0.084	0.141	0.499	131.44	1.785	2	0.409594
36	1.101	0.124	0.425	133.36	2.371	2	0.30565
37	-0.626	0.149	0.346	130.48	0.653	2	0.721283
38	-0.021	0.106	-0.317	130.48	2.642	2	0.266898
39	0.207	0.119	0.228	131.44	6.316	2	0.042511
40	-0.22	0.13	0.645	132.4	0.918	2	0.632007
41	0.472	0.129	-0.019	133.36	3.671	2	0.159498
42	1.17	0.15	-0.017	130.48	10.434	2	0.005423
43	-0.325	0.125	-0.071	130.48	9.026	2	0.010964
44	-0.464	0.134	0.088	133.36	8.746	2	0.012611
45	-0.357	0.138	0.972	126.64	11.043	2	0.004002
46	-0.357	0.151	-0.251	132.4	9.795	2	0.007466
47	-0.856	0.127	0.467	121.84	2.377	2	0.304619
48	-0.688	0.115	0.101	130.48	2.093	2	0.351213
49	0.547	0.123	0.552	133.36	1.486	2	0.47578
50	-0.716	0.119	0.565	134.32	2.539	2	0.280934
51	-0.691	0.124	0.947	134.32	7.935	2	0.018918
52	-0.686	0.126	0.434	133.36	1.06	2	0.588565
53	-1.923	0.157	0.245	134.32	1.956	2	0.376038
54	1.334	0.168	-0.176	122.8	6.878	2	0.032105
55	0.536	0.099	-0.175	122.8	0.362	2	0.834242
56	4.523	0.203	0.176	120.88	0.933	2	0.627103
57	1.948	0.206	0.177	121.84	2.854	2	0.240085
58	0.24	0.123	0.376	126.64	2.039	2	0.360788
59	0.853	0.104	-0.404	126.64	5.766	2	0.055966
60	1.788	0.142	-0.279	130.48	8.902	2	0.01167
61	1.144	0.119	0.176	129.52	1.965	2	0.374382
62	1.789	0.144	-0.034	130.48	7.345	2	0.025416
63	1.788	0.137	-0.3	128.56	6.844	2	0.032644
64	2.991	0.163	-0.123	122.8	1.604	2	0.448521
65	2.908	0.156	-0.307	121.84	7.104	2	0.028668

The questionnaires formed a fair measure of the trait *reasoned risk-taking in decision-making* for this group of principals. The difficulty level estimates of the items ranged from –3.901 to 4.523, whereas the attitude level estimates for the principals ranged from –2.71 to 1.201. For most accurate measurement of persons on a trait, item difficulty should be matched as closely as possible to the person's attitude levels as the standard errors of measurement are least in this case. All of the items with difficulty levels above 1.201 referred to behaviours. Andrich and Styles (1994) have found that behavioural items tend to be more difficult to agree to than attitudinal items. In their paper they argue that attitude and behaviour statements fall on the same measurement continuum with behaviour statements at the higher end as they are harder items to agree to. Thirteen of the 17 behavioural items on the questionnaire had difficulty locations above zero, the mean of the measurement scale. Figure 2 shows the distribution of the attitudinal and behavioural

items and also the location of principals on the measurement scale. The wording of the response sets in the questionnaire was slightly different for attitudinal and behavioural items to clarify the meaning for respondents. However, this may have introduced a limitation in making direct comparisons between the two item sets. Waugh (2005, 2010) indicates that the same items and categories should be used to measure both attitude and behaviour.



Persons **Item Locations**

Figure 2: Item Map showing the person/item distribution with the items identified

The variance among attitude estimates relative to the error variance for each person was acceptable. The separation index, an index similar in principle to the traditional reliability index (Andrich, 1982) can be calculated as the ratio of the estimated true variance of ability relative to the observed variance. The value in this case was 0.633 which indicates the power for test-of-fit is "reasonable". Analogous to the traditional reliability index, the greater the variation in person attitudes the greater the opportunity for the ordering of persons to reveal itself. As the variation in persons' attitudes becomes extreme the separation index tends to one and as variation in persons tends to zero the index of separation tends to zero also. A separation index close to zero would indicate that the differences among the attitude estimates of principals were no greater than would be expected by chance relative to the error of measurement. Re-analysis of the data with items with disordered thresholds and the item with misfitting statistics removed, did not improve fit with the model. The separation index was 0.599 and the power for test-of-fit "reasonable". Andrich, Sheridan & Luo (2005, p.35) assert that this is a common problem when discarding items as precision of measurement is reduced when items are eliminated.

In order to check for any differential performance due to gender, age, level of education, substantive appointment as principal or school location, the responses to each item were divided into categories in each case, thus creating multiple items, one for each response for each item. For example, for level of education there would be four items corresponding to each of the education levels identified in the questionnaire; bachelor, post graduate diploma, Masters and Doctorate. The frame of reference for checking any bias for subscales for each construct, and for individual items is the overall estimate of a single latent trait, reasoned risk-taking in decision-making, for each person, based on a score from items from all subscales. If there is no bias in an item relative to this frame of reference, then the expected value curve for each group, such as principals from each school type, will be equivalent. Otherwise, for chosen values of the latent trait, the expected values will be different.

Item characteristic curves, which are expected value curves, were therefore drawn for each item for each gender, age, level of education, substantive appointment as principal or school location, to determine whether bias was evident. A statistical test was then conducted to determine if any difference observed in the item characteristic curves was significant. Ironson (1983, p.155) indicates that the use of item response theory provides two advantages over traditional methods when attempting to measure item bias. Firstly, it provides a common scale of reference for person ability for each group and for item difficulty which avoids having to make judgements of bias relative to a separate criterion which itself may or may not also have inherent bias. In the context of this analysis, an assumption is made that the bias in the questionnaire as a whole is less than that present in individual items. Secondly, it circumvents problems associated with different distributions in opinion in the two groups of interest as the item response theory models are designed to be insensitive to the different shapes of the distributions.

The differential item functioning analysis found no bias in any item of the questionnaire in relation to gender, age, level of education or school location. Item 49 was found to show bias in relation to substantive appointment as principal. This item asked whether principals referred to training or professional development in their decision-making. Non-substantive principals were significantly more likely to agree that they refer to training or professional development in making decisions. Item 27 was found to show bias in relation to both time employed as a principal and to time employed as a teacher/school administrator. This item asked if the respondent had a great deal of experience in making decisions as a principal. Those principals with less time in the job were more likely to indicate that they had less experience in decision-making. In both cases where differential item functioning was found, the correlation between the demographic factor and the item was intuitively obvious and there was deemed no need to scrutinise the items for any further explanation.

A further Rasch analysis was conducted looking at the subscales determined by each of the constructs in the hypothesised model. The five aspects were: compliance governance mechanism; educative governance mechanism; experience of principals; stakeholder characteristics; and reasoned risk-taking in decision-making. The first construct, compliance mechansim, is measured by 15 items on the questionnaire. The second, educative mechanism, is measured by 5 items. The third, experience, which is proposed to be a mediating variable in the model, is measured by 16 items. The fourth, stakeholder, characteristics, is measured by 13 items. And finally, the dependent variable of reasoned risk-taking in decision-making is measured by 14 items in the questionnaire. Each of these constructs was analysed to determine whether the items included on the questionnaire for each of the constructs formed a scale of measurement for the construct. In RUMM this is achieved through running the creating subtests procedure. This procedure employs a technique similar to that used for deleting items, but “the original items are now regrouped rather than being deleted completely from the analysis” (Andrich, Sheridan & Luo, 2005, *Extending the RUMM2020 Analysis Manual,* p.63*).*

This analysis showed poor fit to the model with a separation index of only 0.34. This outcome could be due to non-linearity of the constructs. It is consistent with the grouped items operating in a non-linear way due to the constructs providing multiple traits for measurement. The low Chi Square (p=0.0000) for the item-trait interaction of the questionnaire as a whole also supports the inference that the scale was not measuring a uni-dimensional trait but was more likely measuring a dominant trait comprised of several dimensions. This is a limitation arising from the analysis and further poor fitting items would need to be deleted to create a unidimensional scale. Rasch item-trait interaction is considered a better method than factor analysis for determining unidimensionality (Smith, 1996; Waugh & Chapman, 2005). However, factor analysis is considered a useful tool for data reduction when multivariate techniques are to be used sequentially (Hair et al., 2006). An exploratory factor analysis was conducted as the next phase of analysis to test this possibility and to confirm the grouping of items within the constructs identified in the hypothesised model.

The rigorous statistical procedures employed in the data analysis resulted in data for 12 items in the questionnaire being discarded from analysis using structural equation modeling. However, the data retained complied with the stringent measurement criteria applied providing confidence in the constructs measured by these data in the hypothesised model. Cavanagh & Romanoski (Waugh, 2005, p.77) indicate that reduction of measurement items through the use of statistical and conceptual procedures during an empirical investigation is acceptable and improves the measurement properties of the survey instrument. They discuss a study of school classroom learning culture where two thirds of the items from their original survey were discarded following Rasch analysis leaving a scale of logical elements for the constructs of interest. Cavanagh & Romanoski then used structural equation modeling to examine the interaction between the elements of the statistically validated structural model.

Assessment of the Structural Model

The analysis discussed above provided assurance that the measurement model was valid and reliable to a degree that would allow confidence in using the measures to test the structural model. This part of the analysis examines the relationships among the constructs to determine whether the measured data support the structural model hypothesised.

The PLS technique utilises a jackknife or a bootstrap procedure to test the significance of parameter estimates (Chin, 2001, p.14; 1998, pp.318-320). The goodness of fit of the model was tested using the root mean square error of approximation. This essentially provides a measure of the predictive power of the model in that it demonstrates the how much of the variance in the construct is explained by the model (Barclay, 1995, p.299). The hypothesised model is useful if it has predictive validity in that the exogenous variables predict the endogenous variable, *reasoned risk-taking in decision-making.* The results shown in Table 4 below indicate support for the predictive validity of the constructs, *Experience*, *Compliance Governance Mechanism*, and *Stakeholder Characteristics*, due to the significant association found with *reasoned risk-taking in decision-making.* However, the construct of *Educative Governance Mechanism* failed to show a significant correlation and therefore is considered to have failed the predictive validity test. This means that this factor failed to predict *reasoned risk-taking in decision-making.*

Table 4: Results for model with no interactions

Hypothesis	Standardised Path Coefficient	t-value
Compliance → Reasoned Risk-taking in decision-making	-0.323	3.55**
Educative → Reasoned Risk-taking in decision-making	-0.004	0.05
Experience → Reasoned Risk-taking in decision-making	0.384	4.54**
Stakeholder Characteristics → Reasoned Risk-taking in decision-making	0.177	2.43*

** $p<0.005$; * $p<0.01$
R^2 for Reasoned Risk-taking in decision-making = 0.477

Table 4 above shows the results of structural analysis for the model with no interactions. Hypotheses 1, 2 and 5 are supported. Hypothesis 3 is not supported. The model explains 47.4% of the variance in reasoned risk-taking in decision-making and thus can be deemed an adequate. Quaddus (2005, p.6) indicates that a value of 25% or more indicates a model with adequate merit.

Table 5: Results for model with interactions

Hypothesis	Standardised Path Coefficient	t-value
Compliance → Reasoned Risk-taking in decision-making	-0.223	2.08**
Educative → Reasoned Risk-taking in decision-making	-0.055	0.55
Experience → Reasoned Risk-taking in decision-making	0.274	2.38*
ComplianceXExperience → Reasoned Risk-taking in decision-making	0.15	1.19
EducativeXExperience → Reasoned Risk-taking in decision-making	0.029	0.32
Stakeholder Characteristics → Reasoned Risk-taking in decision-making	0.164	2.27*

** $p<0.005$; * $p<0.025$
R^2 for Reasoned Risk-taking in decision-making = 0.483

Table 5 above shows the results of structural analysis for the model with multiplicative interaction items. Hypotheses 1, 2 and 5 are supported. Hypotheses 3, 4a and 4b are not supported. The model explains 47.4% of the variance in reasoned risk-taking in decision-making and thus can be deemed an adequate.

Multi-Group Modeling

Multi-group modeling was conducted on the model without interactions, to test whether the model showed differences in applicability when divided into groups based on measures of experience. This provides an alternative method for looking at the moderating effect of the construct *Experience*. This approach tests the moderating effect of *Experience* on the model as a whole whereas the previous interactive analysis looked at the moderating effect on specific paths in the model.

Chin (2004) indicates that this can be approached by taking the standard errors for the structural paths provided by PLS-Graph in the bootstrap output and calculating the t-test for the difference in paths between groups. He recommends use of the Smith-Satterthwaite test if the variance of the samples are assumed different, as is the case in this example.

$$t = \frac{Path_{sample\_1} - Path_{sample\_2}}{\sqrt{SE^2_{sample\_1} + SE^2_{sample\_2}}} \quad \text{with m+n-2 degrees of freedom.}$$

Prior analyses were reviewed to determine measures of the construct *Experience* that were most suitable for a multi-group analysis. Item 4 "*Do you hold the role of principal substantively*" and item 5 "*How long have you been employed in the role of principal*" are the two demographic items measuring experience that were robust enough to remain through each stage of the analysis. They showed sound measurement properties in the Rasch analysis, loaded strongly to the construct experience in the factor analysis and met the requirements of reliability, consistency and validity for the structured equation measurement model.

Review of the descriptive statistics for each of these items showed that item 4 was not suitable for multi-group analysis as the number of principals in each of the two groups, substantive principals and acting principals was too skewed, with only 13% of principals responding being in acting positions. Item 5, however, showed a spread of responses across the categories. Regrouping item 5 "time as principal" into two categories; <5 years and ≥5 years provided two groups with 40% and 60% of responses respectively.

In addition, a multi-group analysis by item 9 "school type" was conducted. Item 9 showed sound measurement properties in the Rasch analysis, but was deleted as a measure of experience in the factor analysis as it was measuring type of experience and did not correlate highly with the other items. This demographic item was often referred to by principals in the preliminary interviews as being significant in whether to take risks in decision-making. It was determined that this provided a reasonable case for looking at principals in different types of schools. The original categories were regrouped into primary and district high schools, with 66% of responses, and secondary schools including agricultural and senior colleges, with 34%.

Multi-group Model by Item 5 – Time as Principal

The multi-group analysis by item 5 was unsuccessful. For the group ≥5 years a PLS processing error occurred as the covariance of the indicators was not positive definite. As a consequence the results of the analysis must be considered invalid and have not been included for further consideration.

Multi-group Model by Item 9 – School Type

Table 6: Structural Model

	Primary and DHS (n=89)		**Secondary** (n=47)		
Hypothesis	Standardised Path Coefficient	t-value	Standardised Path Coefficient	t-value	Multi group t-value (df=134)
Compliance → Reasoned Risk-taking in decision-making	-0.394	4.82***	-0.256	1.53*	0.74
Educative → Reasoned Risk-taking in decision-making	0.108	1.26	-0.275	1.77**	2.31**
Experience → Reasoned Risk-taking in decision-making	0.431	4.77***	0.158	1.64*	1.55*
Stakeholder Characteristics → Reasoned Risk-taking in decision-making	0.251	2.82***	0.335	1.91**	0.43
R^2 for Reasoned Risk-taking in decision-making	0.521		0.598		

*** p<0.005; p<0.05**; p<0.10*

Table 6 above shows the results of structural analysis. Hypotheses 1, 2, 3 and 5 are supported in both the primary and secondary school groups. The model explains 52.1% and 59.8% of the variance in reasoned risk-taking in decision-making in the primary and secondary groups respectively and thus can be deemed an adequate model for each group. Multi-group analysis was conducted for paths that were found to be significant for at least one group. The multi-group t-values in the last column were found using the Smith-Satterthwaite test. Significant differences were found between the different types of schools for the structural paths of *Educative Compliance*

Mechanism on *Reasoned Risk-taking in Decision-making* and for *Experience* on *Reasoned Risk-taking in Decision-making.*

CONCLUSIONS

The final section of this paper discusses the findings and the support found for the hypothesised model. Overall, hypotheses 1, 2 and 5 were supported by all sections of the analysis. Both the model without interactions and the model with multiplicative interaction items showed significant evidence to support these hypotheses. Hypothesis 3 however, was not supported by the model with or without interactions. Hypotheses 4a and 4b were only supported by the multi-group analysis for a sub-group of principals.

Hypothesis 1: More experienced principals will tend to engage in risk-taking behaviour more frequently than new or acting principals.

Hypothesis 1 was supported by all sections of the analysis. Both the model without interactions and the model with multiplicative interaction items showed significant evidence to support this hypothesis. In addition, each part of the multi-group analysis supported this hypothesis also.

The analysis of both of the models, with and without interactions, showed that the strongest structural path within the models was between *Experience* and *Reasoned Risk-taking in Decision-making*. The hypothesis is supported with a finding of significant association. This aligns with the findings from the interviews with principals conducted in the preliminary qualitative phase of the research where experienced principals more often indicated that they preferred greater flexibility to make decisions at the school level to meet outcomes that took account of local circumstances. They expressed a preference for minimal mandatory policy and procedures as they were of the view that their professional expertise would provide sufficient basis for best achieving required outcomes (Trimmer, 2003, p.34).

This association between *Experience* and *Reasoned Risk-taking in Decision-making* was consistent across principals from all school types in the multi-group analysis, but was greater for primary and district high schools. The multi-group analysis divided by school type showed that greater experience influenced risk-taking in decision-making more for primary and DHS principals. However, the difference between groups was not significant.

In contrast, in the preliminary qualitative interviews, secondary school principals more often indicated that they preferred flexibility in decision-making and expressed the view that they had the capacity to make the most appropriate decisions to meet required outcomes in their schools. A possible explanation that accounts for both of these findings is that secondary principals of all levels of experience are more likely to engage in risk-taking in decision-making. The effect of experience is still apparent for secondary principals but may be more pronounced for primary and district high school principals if they are less likely to take risks.

This explanation could be tested through follow-up research that looked more closely at the differences in risk-taking in decision-making in different school types. A subsequent alternative hypothesis for future research could be: Principals with secondary school experience tend to engage in risk-taking behaviour more frequently than primary principals.

Hypothesis 1 was also supported for both agree and disagree groups of principals with greater experience as measured by agreement to items 30 and 31. Within each group there was support for hypothesis 1, with principals more likely to make risk related decisions where they had greater experience. However, whilst the standardised path coefficients were larger for the disagree groups in each case, there was no significant difference found between the groups of principals.

Hypothesis 2: Where school principals interpret the governance mechanism of the regulatory framework as a compliance mechanism there will be a negative relationship to reasoned risk-taking.

Hypothesis 2 was supported by all sections of the analysis. Both the model without interactions and the model with multiplicative interaction items showed significant evidence to support this hypothesis with a strong negative path coefficient. An interesting outcome of the analysis was the importance of retaining items that represented each aspect of the construct identified through the component factor analysis. When such items were removed the functioning of the construct changed to a positive correlation as the construct was altered to be measuring something different from the original.

Principals in both primary and secondary schools, with a compliance view of the governance mechanism, were less likely to take reasoned risks in decision-making.

Hypothesis 3: Where school principals interpret the governance mechanism of the regulatory framework as an educative mechanism there will be a positive relationship to reasoned risk-taking.

Hypothesis 3 was not supported in the analysis of either model, with or without interactions. However, the multi-group modelling did show significant paths for the effect of *Educative Governance Mechanism* on *Reasoned Risk-taking in Decision-making* for some groups of principals. A significant negative effect was found for secondary principals. This result is contrary to what was anticipated. These results are discussed under Hypothesis 4 below.

Hypothesis 4a: There will be an interaction effect between perception of the regulatory framework and experience such that the relationship between the compliance mechanism and reasoned risk-taking is moderated negatively by the experience of principals.

Hypothesis 4b: There will be an interaction effect between perception of the regulatory framework and experience such that the relationship between the educative mechanism and reasoned risk-taking is moderated positively by the experience of principals.

Hypotheses 4a and 4b were not supported by the analysis of the model with interaction items. However, there were mixed results for the multi-group analysis that provide some support for the hypotheses.

Consistent with results of the multi-group analysis divided by school type for hypothesis 1, effect size for principals in primary and district high schools who may be more averse to risk taking in decision-making was greater. There was a negative relationship between compliance view and risk taking for principals of all school types that was greater for principals in primary and district high schools. However, this difference was not significant. This provides support for conduct of additional investigation as described in regards to hypothesis 1, but does not provide sufficient evidence to support hypothesis 4a.

Principals in secondary schools with an educative view were less likely to engage in risk-taking in decision-making, whereas principals in primary schools with an educative view of governance mechanisms were more likely to do so. This supports hypothesis 4b for primary and DHS principals only. This finding was unexpected and does not support hypothesis 4b for secondary school principals. The relationship between secondary principals' views and risk taking in decision-making does not have the same impact as for other principal types. The tendency for secondary principals to take greater risks in decision-making regardless of their view of the governance mechanism appears to dominate. This means that even those secondary principals with an educative view are significantly more likely to take greater risks in decision-making.

Hypothesis 5: Principals of schools with a high degree of uniqueness in the characteristics of key stakeholders within the communities will be more likely to make decisions involving reasoned risk-taking.

Hypothesis 5 was supported by all sections of the analysis. Both the model without interactions and the model with multiplicative interaction items showed significant evidence to support this hypothesis.

Principals in all types of schools with a high degree of uniqueness were more likely to make decisions involving reasoned risk-taking but there was no significant difference between primary and secondary schools. This supports hypothesis 5. Similarly, in the multi-group analysis by items 30 and 31, principals in schools with a high degree of uniqueness were more likely to make decisions involving reasoned risk-taking regardless of their level of experience, providing support for hypothesis 5.

Further Research

The construct experience did not show significant interaction with school type in the multi-group modeling. However, there was some qualitative evidence that school type should be treated independently as a separate construct. In addition, the results showed that having greater experience influenced risk-taking in decision-making more for primary and DHS principals. The hypothesised model could be revised to include school type as a separate construct. School type was included as one of the formative aspects of the construct experience. The factor analysis identified "type of experience" as one of six component factors making up this construct. However, four other factors loaded more highly on the construct and as a consequence the items measuring "type of experience" were dropped from the final analysis. Given the results of the multi-group modeling by school type, there is potential for a further research study to be conducted that distinguishes between school type and experience as separate constructs.

The results of the multi-group analysis suggest that the other constructs in the model operate differently for different school types. Significant differences between all of the structural paths were found between the different types of schools using the Smith-Satterthwaite test. The probability of engaging in risk-taking in decision-making was found to be significantly greater with secondary principals. Principals in primary and district high schools with a compliance view of the governance mechanism were more likely not to take reasoned risks in decision-making. Where principals held an educative view of governance there were also differences found related to school type. Principals in secondary schools with an educative view were less likely to engage in risk-taking in decision-making, whereas principals in primary schools with an educative view of governance mechanisms were more likely to do so. School type may also interact with the construct of stakeholder characteristics with principals in secondary schools with a high degree of uniqueness being more likely to make decisions involving reasoned risk-taking than those in primary schools. A deeper understanding of the involvement and influence of stakeholders may be gained by utilising the approach by Panova (2008, pp. 89-90) where levels of stakeholder participation and influence on higher education institution policies were measured for identified stakeholder groups and policies.

Another potential research opportunity will arise with the move by the Department of Education and Training in Western Australia to create independent public schools that are able to govern themselves autonomously. Decision-making in these schools would no longer be governed by the regulatory framework but by the principal under the authority of a school board. This would create a governance structure similar to independent schools but with a population of students similar to that attending other government schools. A study looking at risk-taking in decision-making in government schools with such an autonomous governance structure would provide a new perspective that has not been able to be considered in the government school context in Western Australia in the past.

In this paper the data collected fit the hypothesised model to confirm a number of the hypotheses. These results assist in improving understanding of factors that impact on risk-taking in decision-making.

REFERENCES

Andrich, David. 1982. An extension of the Rasch model for ratings providing both location and dispersion parameters. *Psychometrika* 47 (1):105-113.

———. 1988. *Rasch Models for Measurement*. Vol. 07-001, *Sage University Papers Series on Quantitative Applications in the Social Sciences*. Beverley Hill: Sage Publications.

———. 1989. Distinctions between assumptions and requirements in measurement in the social sciences. In *Mathematical and Theoretical Systems*, edited by J. A. Keats, R. Taft, R. A. Heath and S. A. Lovebond. B V North Holland: Elsevier Scinece Publishers.

Andrich, David, Sheridan, Barry and G Luo. 2005. *Rasch Unidimensional Measurement Model (RUMM2020):a windows based program using Rasch measurement.* RUMM Laboratory, Perth.

Andrich, David, Barry Sheridan, and A Lyne. 1991. *ASCORE Manual of Procedures*.

Andrich, David, and Irene Styles. 1994. The structural relationship between attitude and behaviour statements from the unfolding perspective. Paper read at 23 International Congress of Applied Psychology, at Madrid, Spain.

Barclay, Donald W, Ronald Thompson, and Christopher A Higgins. 1995. The Partial Least Squares (PLS) Approach to Causal Modeling, Personal Computer Adoption and Use as an Illustration. *Technology Studies: Special Issue on Research Methodology* 2 (2):285-324.

Bennis, W, and B Nanus. 1985. *Leaders: The Strategies for Taking Charge*. New York: Harper and Row.

Bradbury, Keryn. 2009. Govt heralds 'new era' for public education. *ABC News*, 12/08/2009.

Caldwell, Brian J. 1990. School-based Decision-making and Management: International Developments. In *School-based Decision-making and Management*, edited by J. Chapman. London: The Falmer Press.

———. 2006. *Re-imagining Educational Leadership*. Camberwell, Victoria: ACER Press.

Carpenter, Mason A, Timothy G Pollock, and Myleen M Leary. 2003. Testing a Model of Reasoned Risk-Taking: Governance, the Experience of Principals and Agents, and Global Strategy in High-Technology IPO Firms. *Strategic Management Journal* 24:803-820.

Cavanagh, Robert F, and Joseph T Romanoski. 2005. Linking Student Behaviours and Attitudes Towards Information and Communication Technology with Learning Processes, Teacher Instruction and Classroom Environment. In *Frontiers in Educational Psychology*, edited by R. F. Waugh: Nova Science Publishers, Inc.

———. 2005. Sequential Application of Rasch Analysis and Structural Equation Modeling to Investigate Elementary School Classroom Learning Culture. In *Frontiers in Educational Psychology*, edited by R. F. Waugh: Nova Science Publishers, Inc.

Cavanagh, Robert F, and Russell F Waugh. 2004. Secondary School Renewal: The Effect of Classroom Learning Culture on Educational Outcomes. *Learning Environments Research* 7:245-269.

Chin, Wynne W. 1998. The partial least squares approach to structural equation modeling. In Modern Methods for Business Research, edited by G. A. Marcoulides. New Jersey: "Lawrence Erlbaum Associates.

———. 2001. PLS Graph User's Guide 3.0.

———. *Multi-Group analysis with PLS* 2004 [cited 5/2/2008. Available from http://disc-nt.cba.uh.edu/chin/plsfaq/multigroup.htm.

Department of Education and Training. 2009. Unlock Your School's Future, edited by Department of Education and Training. Perth.

Eacott, Scott. 2009. New Look Leaders or a New Look at Leadership. In *Australian Association for Research in Education*. Canberra.

Fidler, Brian. 2001. A Structural Critique of School Effectiveness and School Improvement. In *School Effectiveness and School Improvement*, edited by A. Harris and N. Bennett. London: Continuum.

Fullan, Michael. 1993. *Change Forces Probing the Depths of Educational Reform*. London: The Falmer Press.
Hair, Joseph F Jr, William C Black, Barry J Babin, Rolph E Anderson, and Ronald L Tatham. 2006. Multivariate Data Analysis. Sixth ed. New Jersey: Pearson Education International.
Hoskisson, Robert E, Michael A Hitt, William P Wan, and Daphne Yiu. 1999. Theory and research in strategic management: Swings of a pendulum. *Journal of Management* 25 (3):417-456.
Ironson, G H. 1983. Using Item Response Theory to Measure Bias. In *Applications of item response theory*, edited by R. K. Hambleton. Vancouver: Educational Research Institute of British Coumbia.
Kahneman, Daniel, and Amos Tversky. 1979. Prospect Theory: An analysis of decisions under risk. *Econometrica* 47 (2):263-292.
Karmel, Peter, Jean Blackburn, Greg Hancock, Edward T Jackson, A W Jones, F M Martin, Peter Tannock, M E Thomas, Alice Whitley, Wilfred A White, and J J Willson. 1973. Schools in Australia: Report of the Interim Committee for the Australian Schools Commission. Canberra.
Lunz, Mary E, and John M Linacre. 1998. Measurement Designs Using Multifacet Rasch Modeling. In *Modern Methods for Business Research*, edited by G. A. Marcoulides. New Jersey: Lawrence Erlbaum Associates.
O'Keeffe, Darragh. 2009. Public Schools Get Autonomy. *Education Review* (August):4.
Organisation for Economic Co-operation and Development. 2000. Reducing the risk of policy failure: Challenges for regulatory compliance: OECD.
———. 2004. : OECD.
Panova, A. 2008. Governance Structures and Decision Making in Russian Higher Education Institutions. *Russian Social Science Review* 49 (5):76-93.
Quaddus, Mohammed, and Didi Achjari. 2005. A model for electronic commerce success. *Telecommunications Policy* 29:127-152.
Rasch, G. 1960/1980. *Probabilistic models for some intelligence and attainment tests*. Chicago: University of Chicago Press.
Rumelt, Richard P, Dan Schendel, and David J Teece. 1991. Strategic Management and Economics. *Strategic Management Journal* 12:5-29.
Sarason, S. 1982. *The Culture of School and the Problem of Change*. Revised ed. Needham Heights: Allyn and Bacon.
School Education Act. 1999. Parliament of Western Australia.
Seddon, Terri, Lawrence Angus, and Millicent E Poole. 1990. Pressures on the Move to School-based Decision-making and Management. In *School-based Decision-making and Management*, edited by J. Chapman. London: The Falmer Press.
Sergiovanni, T J. 2000. *The lifeworld of leadership: creating culture, community and personal meaning in our schools*. San Francisco: Jossey-Bass Publishers.
Sheridan, Barry. 2007. Home Page Feedback. Perth, 13/08/2007.
Smith, R.M. 1996. A comparison of methods for determining dimensionality in Rasch measurement. In *Structural Equation Modeling,* 3 (1), 25-40.
SPSS Version 12. 2003.Apache Software Foundation.
Starr, Karen. 2008. Whose Risk? 'Managing' Risk in the Principalship. In *Australian Association for Research in Education*. Brisbane.
Trimmer, Karen. 2003a. Review of the Regulatory Framework: Education Department of Western Australia.
———. 2003b. Review of the West Australian Department of Education Regulatory Framework. Paper read at Australian Association of Research in Education Conference, at Auckland, New Zealand.
Waugh, Russell F, 2010. Specialized Rasch measures applied at the forefront of education. New York: Nova Science Publishers, Inc.
Waugh, Russell F, ed. 2005. Frontiers in Educational Psychology. New York: Nova Science Publishers, Inc.

Waugh, Russell, and Chapman, E.S. 2005. An analysis of dimensionality using factor analysis (True Score Theory) and Rasch Measurement: What is the difference? Which method is better? In *Journal of Applied Measurement,* 6 (1), 80-99.

Wiseman, Robert M, and Luis R Gomez-Mejia. 1998. A behavioral agency model of managerial risk taking. *Academy of Management Review* 23 (1):133-153.

Wiseman, Robert M, Luis R Gomez-Mejia, and M Fugate. 2000. Rethinking compensation risk. In *Compensation in Organizations*, edited by S. L. Rynes and B. Gerhart. San Francisco: Josey-Bass.