MALTA REVIEW OF EDUCATIONAL RESEARCH

Volume 9, No.2, pp. 331-353
Faculty of Education©, UoM, 2015

The Influence of Teacher Behaviours on Pupils' Mathematical Attainment at Age 6

Lara Said
Faculty of Education, University of Malta
lara.said@um.edu.mt

Abstract: Teachers' instructional behaviours are proximal to pupil learning but not isolated from the broader setting of education. The overall aim of this paper is to explore the influence of teaching on pupil attainment. Utilising a large national sample of pupils' standardised outcomes, this paper revisits and reanalyses data from a 2005 study called 'Mathematics in Maltese Primary Schools' (MIMPS). The study employed random stratified sampling methods to sample pupils (n = 1,628), in Year 2 classrooms (n = 89) in primary schools (n = 41). Pupils were administered Maths 6 (NFER). Results from multilevel analyses reveal, that after adjusting for the contribution of pupil, classroom and school level factors, pupil ability, curriculum coverage, teacher behaviour and head teacher age were elicited as significant and influential predictors of pupil attainment at age 6. The findings highlight the importance of quality teaching and instruction for pupil attainment. The author concludes by recommending the implementation of a system to monitor pupils' baseline and later attainment outcomes in tandem with the contexts and processes associated with classrooms and schools.

Keywords: mathematical attainment, early childhood education, quality teaching

Introduction

Why do some pupils achieve better outcomes in Mathematics than some of their peers? To address this question, the paper examines the association between pupils' mathematical attainment at age 6 and educational factors. The focus of this work is quality teaching, identified by many researchers as of central importance to pupil learning (Askew, Brown, Rhodes, Johnson & Wiliam 1997; Darling-Hammond, 2012; Hattie, 2009; Mujis & Reynolds, 2000).

'Mathematics in Maltese Primary School' or MIMPS was initially inspired by the 'Effective Provision of Pre-school Education' or EPPE (Sammons et al., 2002; Sylva et al., 1999). EPPE examined the influence of pre-school education on the achievement of pupils between the ages of 3 to 7 years in the United Kingdom (UK). Children with multiple disadvantages or children at risk were found to develop at a slower pace than their 'privileged' counterparts. In England, Anders et al. (2010) defined pupils at risk of experiencing learning delay as those pupils with special educational needs who have significantly greater difficulty learning than the majority of children of the same age. Leroy and Symes (2001) include children who fail to learn because of their social circumstances.

Locally, larger-in-scale studies that identify the predictors of pupil attainment have been generally limited to a pupils-in-schools design. The first three local studies based on a school effectiveness design include: a survey of pupil attainment for literacy at age 6 and at age 9 in Maltese and English (Mifsud et al., 2000; Mifsud et al., 2004) and a survey of pupil attainment for numeracy at age 5 (Mifsud, Richardson & Hutchison, 2005). The Numeracy Survey (Mifsud, Richardson & Hutchison, 2005) identified various characteristics that predict pupil attainment at age 5 for mathematics including: sex, time spent at preschool, special needs, father's occupation, mother's occupation, father's education, mother's education, family structure, number of classrooms, type of school and school district associated with 4,666 pupils in 102 primary schools. Whilst the Maltese surveys for literacy and numeracy constitute an important first, they were not designed to examine the contribution of classroom factors on pupil attainment.

Quality of Teaching

Findings from the International Mathematics and Science Studies (TIMSS) by Mullis et al. (2008) and Mullis et al. (2012) compare the mathematics and science outcomes of 14 year-old pupils across different countries, across the globe. Results indicate that the performance of Maltese students in mathematics lags significantly below the international 500 points average. Results for 2007 to 2011 show that the outcomes, associated with two different cohorts of Maltese pupils, do not vary significantly. Of the 45 countries participating in TIMSS 2011, Malta ranked 28th, with the mean attainment of Maltese pupils for mathematics standing at 496 marks. In 2007, the average attainment of Maltese students for mathematics, was similar - 488 marks. This implies that the local longer-term pattern of pupil attainment for mathematics did not differ significantly over a five-year period. It also suggests that educational provision for mathematics did not differ significantly in quality.

Findings from TIMSS (2011) reveal that engaging mathematical instruction is central to pupil achievement. The importance of quality of teacher instruction has also been repeatedly evidenced in other large-scale studies of pupil achievement such as the International School Effectiveness Research Project (ISERP) (Reynolds, et al., 2002), EPPE (Sylva et al., 1999), the now classic Junior Schools Project (Mortimore et al., 1988), the Effective Teachers of Numeracy Study (Askew et al., 1997), the examination of pupil achievement in primary schools in Cyprus for mathematics (Kyriakides et al., 2000; Campbell et al., 2004) and the Evaluation of the Mathematics Enhancement Project (Primary), also known as GATSBY, by Mujis & Reynolds (2000).

Quality of teaching and instruction is known to differ significantly, even amongst teachers within the same school (Berliner, 1985). Processes such as coverage of the curriculum (Marzano, 2003) and effective teacher behaviours (Mujis & Reynolds, 2000) are amongst the teaching factors associated with effective classroom processes (Ko & Sammons, 2010) and effective practice (Day et al., 2006). Teachers are central to life in the classroom (Darling-Hammond, 2012). In catering for the diverse learning needs of pupils, teachers must consider all pupils as capable of learning (Duncan, 2007). Teachers are key for the learning needs of at-risk pupils (Anders et al., 2010). Teachers matter (Day et al., 2007) and teaching makes a positive difference to pupil outcome (Mortimore et al., 1988). Ko & Sammons (2010) describe teachers in effective classrooms in England as: "clear about instructional goals; knowledgeable about curriculum content and the strategies for teaching it" (p.15).

Teachers' instructional behaviours are amongst the more observable elements of teaching (Hattie, 2009; Mujis & Reynolds, 2000) and easier to observe than pupil learning (Day et al., 2006). Effective teachers of mathematics frequently engage in behaviours associated with increased pupil gain (Mujis & Reynolds, 2000). In the UK, Mujis & Reynolds (2001) identified fifty-seven (57) behaviours of effective teachers which they grouped under eight instructional categories: classroom management; classroom behavior; focusing and maintaining attention on the lesson; review and practice; skills in questioning; variety of teaching methods; and a positive classroom climate.

Examining the Influence of Teachers' Instructional Behaviours for Pupil Attainment at Age 6 for Mathematics

The present study asks: 'Which teacher behaviours influence the attainment outcomes of Maltese pupils, for mathematics, at age 6 after adjusting for factors at the pupil, the classroom and the school level?' This question is drawn from MIMPS, a study that utilised the baseline scores attained by the population of Year 1 pupils at age 5 (n = 4,662) who participated in the Numeracy Survey (Mifsud, Richardson & Hutchison, 2005) to quantify pupil

attainment at age 6, pupil progress from age 5 to age 6, as well as qualifying the behaviours and beliefs of Year 2 teachers and head teachers in effective and less effective schools. MIMPS adopted a mixed design to secure increased synergy (Day et al., 2008) and to stray from the dichotomy when singularly applying quantitative or qualitative methods (Brannen, 2005; Creswell, 2009; Tashakkori & Teddlie, 2003). MIMPS was designed to conform to pre-established quality criteria for discriminant 3-level analyses (Goldstein & Spiegelhalter, 1996; Scheerens, 1992) through the application of appropriate multilevel techniques (Raudenbush & Bryk, 2002). In view of the narrower focus of this paper, the Comprehensive Model of Educational Effectiveness by Creemers (1994) was considered more appropriate, than the one by Kyriakides, Creemers and Antoniou (2009).

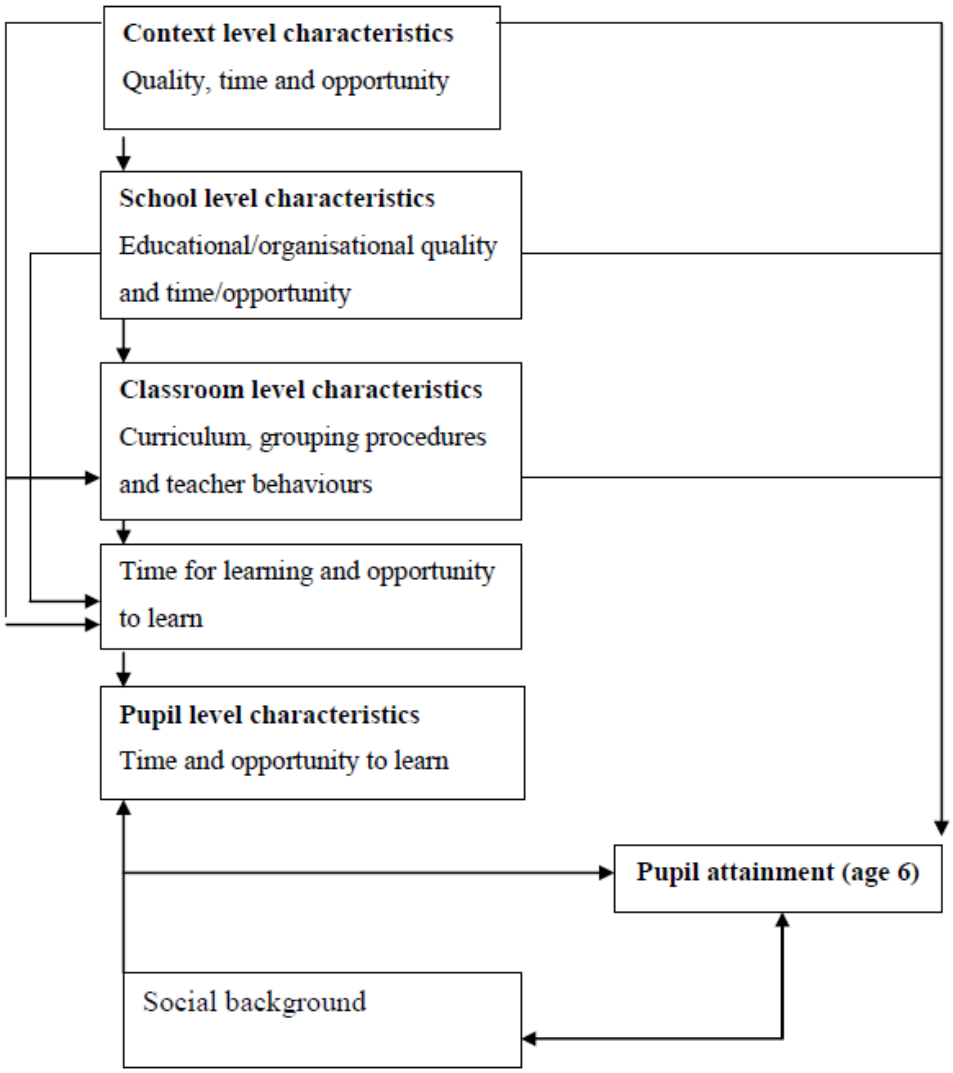

Figure 1 *– A Local-Specific Comprehensive Model for Instructional Effectiveness*
Summarised from Creemers, 1994:119 and adapted by Said (2013)

Variables

Various factors at the pupil, the classroom and the school level are known to contribute differentially to pupil outcome (Mortimore et al., 1988; Creemers, 1994). Pupil ability and background factors, such as the socio-economic status of parents, are associated with significant variations in pupil outcome (Duckworth, 2007). Characteristics such as the status of the family, the home town of pupils, first language, attendance at preschool, private tuition and learning support have also been reported to predict pupil attainment (Mifsud et al., 2000, Mifsud et al., 2004; Mifsud et al., 2005). In Malta, some at-risk pupils are supported by a learning support assistant or by a complementary teacher. Though this factor plays out at the classroom level, such support is known to vary considerably even amongst individual at-risk pupils within the same classroom. Therefore, these variables were included at the pupil level. Similarly, *seating arrangements* was included at the pupil level.

At the classroom level, teaching and instructional factors (Creemers, 1994), particularly those connected with curriculum coverage (Marzano, 2003) and teacher behaviours (Mujis & Reynolds, 2000) are important predictors of pupil outcome. The size of the classroom, socio-economic composition predominant in classrooms, the predominant language of instruction and amount of homework also predict pupil attainment. Teacher attributes, such as *sex*, age, qualifications, *first language* and *experience teaching primary* are also considered as predictors of attainment.

School level predictors of pupil attainment were limited to opportunity factors with regards to the theoretical framework by Creemers (1994), these included: *type* and *size* of school, *school's socio-economic composition*. Said (2013) details the construction of an index based on father's/mother's occupation and father's/mother's education. Qualities of the headteacher were also included. Affiliated variables refer to: sex, age, qualifications, length of time teaching and headteaching.

Research Instruments

Different instrunments were administered to identify the factors influential for pupil. An age-standardised and back-translated version of Maths 6 (NFER) was administered during May 2005 to collate data about pupil outcome. The classroom observation tool, the Mathematics Enhancement Classroom Observation Record (MECORS) by Mujis & Reynolds (2000), collated data about the frequency and quality of behaviours observed of teachers during lessons. Three tailor-made questionnaires, including a parent/guardian, a teacher and a head teacher questionnaire were administered to collate background and/or contextual data. Field notes

referring to school-wide conditions and the daily lives of teachers in classrooms and head teachers in schools were also kept.

Assessing Pupil Outcome: The Maths 6 Test

The attainment outcomes of Maltese pupils at age 6 were assessed on the basis of a UK test. Maths 6 (NFER) which was administered during May 2005 consisted of 26 test items. A year earlier the same cohort of pupils had been tested on Maths 5 (NFER). Pupil recruitment was on the basis of informed consent from parents/guardians. The test took between 30 to 50 minutes to complete and was administered to no more than five pupils at a time by a team of 33 trained researchers Responses to the Maths 6 test were found to be internally reliable on the basis of Cronbach's alpha ($\alpha = 0.81$). This is the same as that reported for England. Said (2006) highlights the relevance of this test by comparing items in Maths 6 with topics in ABACUS.

Back translation methods were applied to translate the test from English to Maltese (Said, 2013). Logistic regression techniques (Kim, 2001; Zumbo, 1999) checked for language bias between the Maltese and the English versions of Maths 6 associated with an achieved sample of 1,937 pupils. On the basis of the differential item classification system by Gierl, Rogers & Klinger (1999) and recommendations by Sireci (1997) a difference of 1.84 standardised marks between the Maltese and the English versions of the Maths 6 test was considered as minimal.

Quality of Primary Teachers' Instructional Behaviours: MECORS

MECORS is a classroom observation tool that was developed by Mujis & Reynolds (2000) in the UK to quantify and qualify the frequency of effective teacher behaviours during lessons of mathematics. Both Part A and Part B were designed to be administered by a trained researcher. Part A involved collating notes about instructional conditions in the classroom. Part B involved the rating of teacher behaviours on a Likert scale ranging from: 1 (never observed), 2 (occasionally observed), 3 (sometimes observed), 4 (frequently observed) and 5 (consistently observed). MECORS was piloted in Malta during May 2004 with a sample of 17 Year 2 teachers located in eight randomly selected pilot study schools; not included as part of the main study. During the main data exercise each of the 89 teachers were observed twice; first in January to February 2005 (round A) and then from March to April 2005 (round B). The same order of teacher observation was maintained during each round. Each teacher was observed by each of the two researchers. Inter-observer reliability, ascertained during a stage preceeding the main data exercise but following the pilot study, was the same as that established by Mujis & Reynolds amongst their four observers at $k = 0.81$, $p < .001$. Said (2013) details how no significant differences in the quality and

frequency observed of the same teachers were elicited during observation rounds A and B.

Contextual and Other Instructional Data: the Survey Questionnaires

Three tailor-made questionnaires, piloted during October to December 2004, were administered to gather broader data. The parent/guardian questionnaire was administered in April 2005 to collate background information about pupil ability, parental occupation and education as well as family structure. The questionnaire was made available to parents in both Maltese and English. Parents were also asked on the basis of informed consent for permission regarding the participation of their child in MIMPS. The teacher survey questionnaire was administered to participating Year 2 teachers during March 2005. Part A required teachers to provide information about their attributes such as age, sex, qualifications and experience. Part B required teachers to answer to 47 belief statements about the teaching and learning of mathematics (Askew et al., 1997).

Sampling

Said (2013) details how she employed a multistage, stratified method of sampling (Goldstein, 1987; Teddlie & Stringfield, 1993) to target different groups of participants for entry into the MIMPS study. First a minimum figure of 368 (95%) pupils/parents and a maximum figure of 1,400 (99%) pupils/parents was calculated according to Yamane (1967). For the purpose of discriminant analysis, classrooms had to exceed 50 (Maas & Hox, 2001) and schools 30 (Kreft, 1996). To cater for attrition, 41 schools, 99 classrooms and approximately 2,000 pupils were targeted for entry. Eventually, data for 1,628 pupils in 89 Year 2 classrooms in 37 primary schools could be matched for multilevel analysis.

Age-Standardisation of Pupil Scores

Age is a discriminating factor. Of the 1,628 pupils, 415 were in the eldest category (January to March), 409 pupils were in the elder category (April to June), 432 pupils were in the younger category (July to September) and 372 pupils were in the youngest category (October to December). As expected, significant differences in pupil attainment were elicited depending on age ($F = 5.200$, $df = 3$, $p < .001$). Since every month increase approximated 0.4 marks, this implied as much as 4.4 marks difference between January and December-born pupils. To statistically control for age, pupils' scores were standardised according to Schagen (1990). The distribution in pupils' age-standardisded scores was found to be within the bounds of normality ($Z = 1.316$ $p = 0.063$) in spite of an apparent ceiling effect (Figure 2).

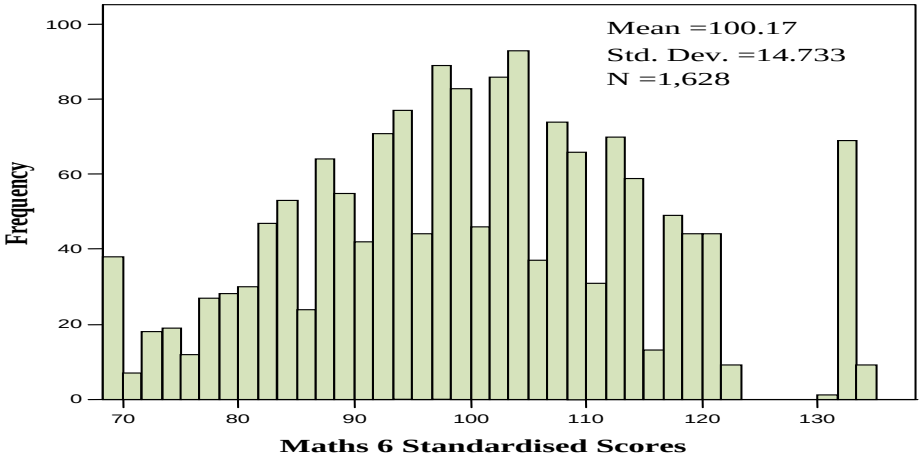

Figure 2 – *Distribution of Age-Standardised Scores*

On the basis of this distribution, pupils who scored between: (1) 69 to 71 marks attained far below average at -2 sd, (2) 71.1 to 85.6 marks attained below average at -1 sd, (3) 99.95 to 114.6 marks did not deviate significantly below or above the average line of attainment, (4) 114.7 to 129.1 marks attained above average at 1 sd, and (5) 129.2 to 141 marks attained far above average at 2 sd.

A Three-Pronged Analytical Strategy

Three forms of quantitative analysis were applied: single-level, structural and multilevel. Single-level analyses descriptive analyses that explore, in this case, the association between pupil attainment and singular factors such as pupil age and class size. These are useful in highlighting the possible issues at stake. However, such analyses are limited and because ignoring nested structures is likely to lead to unreliable standard errors associated with coefficients (Anders et al., 2010). In contrast, multilevel models can simultaneously adjust for relationships within and across levels (Sammons & Smees, 1997). These more complex models are therefore powerful devices for representing socio-educational reality (Goldstein, 1998; Snijders & Bosker, 1999) and necessary for the more accurate treatment of nested forms of multivariate data (Raudenbush & Bryk, 2002). There are also structural equation models that examine the underlying structures associated with contructs (Bryne, 2010) and now considered essential for the validation of instruments (Bryne, 2001).

Variations in Pupils' Age 6 Outcomes for Mathematics (Single-Level Analyses)

Pupils in the Malta sample fared in a comparably similar fashion to pupils in the UK. This occurred for cognitive processes such as: mathematical interpretation (Malta = 85%, UK = 80%), non-numerical processes (Malta = 82%, UK = 77%), computation/knowledge (Malta = 81%, UK = 74%), understanding number (Malta = 70%, UK = 72%) and mathematical application (Malta & UK = 69%). Table 1 lists the percentage of correct responses per Maths 6 test item for Malta.

Table 1 – *Percent Correct of Items in Maths 6*

Maths 6 items	%	Maths 6 items	%
Understanding number	**70.02**	**Mathematical application**	**68.87**
Stories of (7)	75.60	Story sums – sharing (3)	82.80
Ordinal numbers (11)	81.10	Story sums – subtraction (6)	88.90
Stories of (12)	55.10	Patterns (8)	68.60
Between numbers (15)	85.30	Bills (13)	73.20
Value of numbers (20)	90.00	Addition (14)	91.80
Recognition of fractions	81.80	Pairs (16)	32.80
Stories of (22)	82.40	Story Sums - multiplication	43.00
Non-numerical processes	**81.65**	**Mathematical interpretation**	**84.60**
Shapes – properties (4)	75.70	Sets (1)	93.20
Shapes – properties (9)	88.60	Bar graphs – addition (10)	76.00
Size (23)	90.60		
Computation/knowledge	**80.68**		
Shapes (2)	80.20	Subtraction (18)	87.20
Doubles (5)	70.80	Addition with money (19)	69.30
Shapes – recognition (17)	82.10	Clock, hours (26)	93.00
Shapes – properties (24)			

Variations in pupil outcome are partly influenced by socio-economic characteristics such as those affiliated with the occupational and the educational (Duckworth, 2007),. Significant variations in pupil outcome were elcited depending on father's ($F = 4.460$, $df = 6$, $p < .001$), and mother's occupation ($F = 5.200$, $df = 6$, $p < .001$) as well as mother's education ($F = 3.958$, $df = 4$, $p < .001$). Table 2 describes the gap in marks attained between the more and less socio-economically advantaged pupils.

Table 2 – Mean Pupil Outcome by Parental Background

	Father's occupation	**Mean score**	**s.d**	**Difference in marks**
High	Professional (n = 121)	104.00	14.50	Reference
	Managerial (n = 229)	104.00	13.30	No difference
Medium	Higher clerical (n = 325)	100.00	15.20	4.00
	Skilled manual (n = 567)	98.80	14.30	5.20
Low	Semi/un-skilled (n = 184)	100.00	15.08	4.00
	Unemployed (n = 5)	na	na	na
	Other (n = 197)	97.30	15.40	6.70
	Mother's occupation			
High	Professional (n = 78)	106.00	16.00	Reference
	Managerial (n = 65)	106.00	13.10	No difference
Medium	Higher clerical (n = 173)	102.00	14.60	4.00
	Skilled manual (n = 99)	101.00	11.80	5.00
Low	Semi/un-skilled (n = 34)	97.20	14.50	8.80
	Unemployed (n = 1,094)	99.30	14.70	6.70
	Other (n = 85)	96.50	13.20	9.50
	Mother's education			
High	Tertiary (n = 158)	104.00	14.90	Reference
Medium	Secondary (n = 1,035)	99.20	14.50	4.80
	Sixth form (n =329)	102.00	14.40	2.00
Low	No schooling (n = 1)	na	na	na
	Primary (n = 103)	99.00	18.90	5.00

School and Classroom Characteristics (Single-Level Analysis)

Of the 37 schools associated with the matched sample of 1,628 pupils, 24 (64.86%) were from the state sector, 9 (24.32%) from the private church sector and 4 (10.82%) from the private independent school sector. Twenty-two (22, 59.46%) schools were small with 1 to 2 classrooms in each year group. Eleven (11, 29.73%) schools were medium-sized (3 to 4 classrooms) and four sch%) were large (5 to 6 classrooms). Seventeen (17, 45.95%) head teachers were male and 20 (54.05%) female. Five (5, 13.51%) head teachers were aged between 35 to 44 years, 15 (40.54%) head teachers were between 45 to 54 years and 17 (45.95%) head teachers were between 55 to 61 years. The over-riding majority of Year 2 teachers were female (n = 87, 97.75%). Of the 89 classrooms, more than half (56.18%) were medium (16 to 25 pupils) in size, 41.57% were large (26 to 30 pupils) and 2.25% were small (up to 15 pupils. Of the 63 ABACUS topics in the planned curriculum, 68 (76.40%) teachers covered at least 41 topics and 21 (23.59%) teachers covered at least 51 topics.

Time dedicated to lessons of mathematics varied considerably. On average, the school day in state schools is 7 hours in duration in comparison to the 6 hours 20 minute school day in private schools. In state schools lessons of mathematics last for an average of 40 minutes in comparison to the 55 minutes in private schools. On average, typically-developing pupils in state schools experience 175 hours of lesson time whilst typically-developing pupils in private schools experience 243 hours of lesson time. Consequently, pupils in private schools experienced on average 68 hours, or 27%, more in lesson time than pupils in state schools. In state schools pupils with special needs supported by a learning support assistant experienced a different kind of lesson time. This is due to the practice of support assistants acting as "verbal scribes" to explain concepts even during explanations. Therefore, on average special needs pupils with learning support "lose" 160 hours of lesson time as delivered by their class teacher. This is in contrast to the 194 hours of lesson time experienced by pupils with special needs in private schools. Private schools did tend to regulate more the instructional behaviour of learning support assistants. Pupils in state schools considered as experiencing difficulty in learning mathematics "lose" 105 hours of contact time with their teacher during lessons because of small-group support outside of the classroom. In private schools, pupils experiencing difficulty with learning are supported by a complementary teacher during seat-work. On average, such pupils benefited from 200 hours of lesson time.

Validated Local-Specific Behaviours of Year 2 Teachers for Mathematics (Structural Analysis)

Part B of the Mathematics Enhancement Classroom Observation Record (Mujis & Reynolds, 2001) collated data about teacher behaviours. A first-order CFA model sought to validate the eight-category instructional structure. Said (2013) details how she confirmed a local-specific structure comprised of five instructional behaviour factors: *practice, questioning and methods*, *orderly climate*, *management*, *making time* and *broader climate/rewards* (RMSEA = .058, CFI = .968, χ^2 = 308.4, df = 199, p < .001). Table 3 lists mean scores for the local-specific structurally validated set of teachers' instructional behaviours. It is useful to note that mean scores over three indicate that most teachers engaged in the behaviour. Conversely, mean scores under three indicate that most teachers did not engage in this behaviour.

Table 3 – Mean Scores for Observed Teacher Behaviours

Practice, questioning and methods	**Mean**	**s.d**
Presents material clearly (14)	3.83	0.842
Offers assistance to pupils (20)	3.03	1.176
Summarises the lesson (22)	3.18	1.140
Asks academic mathematical questions (26)	3.56	1.131

Probes further when responses are incorrect (28)	2.76	1.248
Appropriate wait-time between questions/responses (32)	4.02	1.073
Notes pupils' mistakes (33)	3.35	1.132
Gives positive academic feedback (38)	3.64	0.916
Uses a variety of explanations that differ in complexity (47)	4.11	0.898
Uses a variety of instructional methods (48)	3.41	0.900
Orderly climate		
Sees that rules and consequences are clearly understood (1)	4.75	0.716
Conveys genuine concern for pupils (54)	3.86	0.841
Displays pupils' work in the classroom (56)	3.01	1.115
Management		
Sees that disruptions are limited (5)	1.83	1.256
Asks pupils for more than one solution (31)	2.59	1.198
Encourages interaction/communication (53)	3.90	0.870
Making time		
Uses time during class transitions effectively (3)	4.02	1.044
Corrects behaviour accurately (8)	4.26	0.676
Guides pupils through errors (34)	4.33	0.900
Broader climate and rewards		
Tasks/materials are collected/distributed effectively (4)	3.56	1.373
Uses a reward system to manage pupil behaviour (6)	3.21	1.690
Knows and uses pupils' names (55)	4.90	0.577

Identifying Factors Significant for Pupil Attainment: (Multilevel Analyses)

Multilevel modelling disentangles the contribution of factors at the pupil, the classroom and the school levels. A null model first examined the association between pupil, classroom and school level factors with pupil attainment at age 6. The intercept (100.794, se 1.464) refers to the mean standardised score achieved by pupils. The variance in the null model was as follows: pupil (163.103, 70%), classroom (6.267, 2.61%) and school (70.771, 29.47%). The next model that was constructed was the pupil model, followed by the classroom models (classroom context and teacher behavior) and then the school model. The pupil level model explained 6.58% of the total variance, of which 6.94% was explained by the pupil level, 0.19% by the classroom level and 0.60% by the school level. The classroom level explained 25.27% of the total variance. Of this variance 8% was attributable to teacher behaviour factors. Of the variance associated with these factors, none of the variance was explained by the pupil level, 1.06% of the variance was explained by the classroom level and 6.93% of the variance was explained by the school level. This implies that

factors at the school level impinge on the quality and frequency of teacher behaviours at the lower level of the classroom.

Various factors were elicited as significant predictors of pupil attainment. At the pupil level, the discriminating influence associated with variations in pupil ability is significant; especially, between typically-developing and at risk pupils (pe = -4.676, se = 1.695, $p < .001$). At the classroom level, results from the same model show the significance of increased curriculum coverage (pe = 8.726, se = 3.403, $p < .05$), five beliefs held by Maltese Year 2 teachers and some teacher behaviours. Behaviours significant for attainment are, teacher: displays pupils' work in the classroom (pe = -7.176, se = 2.608, $p < .01$), sees that disruptions are limited (pe = 3.456, se = 1.555, $p < .05$), prepares an inviting and cheerful classroom (pe = 5.578, se = 1.393, $p < .01$) and uses a reward system to manage pupil behaviour (pe = 1.520, se = 0.577, $p < .05$). At the school level, differences in head teacher age were elicited as significant in their differential contribution towards pupil attainment (pe = -7.174, se = 2.217, $p < .01$).

Other characteristics at the pupil, the classroom and the school level were not elicited as significant predictors of pupil attainment. At the pupil level, these refer to: sex, father's education, parental status, home district, first language, preschool, private lessons and seating arrangements. At the classroom level, these refer to: teacher attributes, broader classroom features and instructional processes. Teacher attributes not significant for pupil attainment include: age, first language, teaching qualifications and experience teaching at primary school. Characteristics associated with the broader classroom context, and not significant for pupil attainment include: predominant socio-economic status associated with parental occupation and education, class size, amount of *homework* set, *lesson duration* and language of instruction. Eighteen (18) of the teacher behaviour variables validated for Malta in Table 3 were also not elicited as significant predictors of pupil attainment. At the school level characteristics not significant for pupil attainment, refer to head teacher attributes. These include: age, first language, teaching qualifications, experience teaching at primary school and experience head teaching.

Influence of Factors Significant for Pupil Attainment

Parameter estimates are statistical devices that express power (Snijders, 2005). The associated influence is expressed in average percentiles for a group in comparison to a reference group. Effect sizes range from 0 (no effect) to ±1, small ($d = .2$), medium ($d = .5$) and large ($d = .8$). Effect sizes in Table 4 were calculated on formulae from Tymms, Merrell and Henderson (1997).

Table 4 – Influential Pupil Level Factors

Pupil level (reference category)	PE	p	SE	Z	Effect size
At risk (typically-developing)	-4.673	***	1.695	-0.754	-0.38
LSA support	-4.015	**	1.015	-0.759	-0.33
Complementary teacher support (typically-developing)	-6.340	***	1.006	-0.643	-0.52
Father's occupation (medium)					
High	1.508	*	0.407	0.302	0.12
Low	-2.540	ns	1.180	-0.238	-0.20
Mother's occupation (medium)					
High	1.424	ns	0.742	0.457	0.15
Low	-1.935	*	0.442	-0.069	-0.16
Mother's education (medium)					
High	2.268	*	0.887	0.147	0.19
Low	-1.291	ns	1.126	-0.039	0.10

na = not applicable (cases amounted to 5 or less), ns = not significant,
*p < 0.05, **p < 0.01, ***p < 0.001

Quality Support Makes a Difference

Differences in pupil ability influence pupils' capacity to learn. The widest gap in marks was elicited on the basis of ability (F = 10.437, df = 1, p < .001). Typically-developing pupils (n =1,359) attained an average of 101 marks (sd = 14.46) in comparison to at-risk pupils (n = 269) who attained an average of 90.01 marks (sd = 15.70. From the at-risk pupil group, 75 pupils had been diagnosed, sometime prior to Year 2, with a condition which negatively influenced their ability to learn. Another 194 pupils had been identified by their Year 1 class teacher as experiencing difficulty in learning mathematics. Coefficients in Table 4 above show that on average at risk pupils achieve 4.6 standardised scores less than their typically-developing peers. Interestingly, the influence of disadvantage is greater for pupils without statements and supported by a complementary teacher than for their at risk peers with statements and supported by a learning support assistant.

The Mixed Compensatory Effects of Schooling

Differences in the socio-economic background of pupils were found to contribute differentially towards pupil outcomes, particularly between pupils with parents in high category occupational/educational backgrounds and pupils with parents with low category occupational/educational

backgrounds. However, the influence of classroom and school level factors also compensate such effects. Compensatory effects are therefore mixed. On the one hand, the gap in attainment between pupils with parents from the high and medium occupational/educational backgrounds remains. On the other hand, schooling mitigates the discriminatory effects of socio-economic background between pupils with parents from the medium and low occupational/educational backgrounds.

Quality Instruction Makes a Difference

Increased curriculum coverage and teachers' instructional behaviours such as when the teacher: *displays pupils' work*, *sees that disruptions are limited*, *prepares and inviting and cheerful classroom* and *uses a reward system to manage pupil behaviour* are influential for pupil attainment.

Table 5 – Instructional Factors Influential on Pupil Attainment

Classroom level (reference category)	**PE**	**p**	**SE**	**Z**	**Effect size**
ABACUS (up to spring)					
Up to summer	8.726	*	3.403	0.101	0.72
Teacher behaviours					
displays pupils work in the classroom (rarely observed)					
Somewhat observed	2.871	*	0.806	0.008	0.24
Frequently observed	4.682	***	1.407	0.102	0.38
sees that disruptions are limited (rarely observed)					
Somewhat observed	na		na	na	na
Frequently observed	3.427	*	1.152	0.015	0.28
prepares an inviting/ cheerful classroom (rarely observed)					
Somewhat observed	-5.326	***	1.201	-0.287	-0.27
Frequently observed	-2.218	***	0.187	-0.147	-0.18
uses a reward system to manage pupil behaviour (6, rarely observed)					
Somewhat observed	-1.235	*	0.526	-0.302	-0.10
Frequently observed	-0.927	*	0.318	-0.100	-0.08

na = not applicable (cases amounted to 5 or less), *p < 0.05, ***p < 0.001

Teachers who covered more than 51 topics exert a considerable positive influence on pupil attainment. The direction of effects associated with behaviours known to enhance the instruction of mathematics were not found to be consistently positive. For example teachers who were *somewhat observed* in *preparing an inviting and cheerful classroom* and teachers who *use a reward system to manage pupil behaviour* are associated with a negative effect on pupil attainment. The implementation of singular behaviours, even if effective, does not suffice to make a difference to pupil attainment. It is clear that pupil attainment is due to an interplay of contextual and process factors.

Younger Headteachers Make a Difference

Younger head teachers exert a more positive effect than older head teachers for pupil attainment. Effect statistics show that headteachers aged between 35 to 44 years exert a greater positive influence (pe = 7.100, $p < .01$, se = 1.427, $d = .58$) than headteachers aged between 45 to 54 years (pe = 3.174, $p < .01$, se = 0.817, $d = .26$) in comparison to headteachers aged 55 to 61 years. The significance of headteacher age as a predictor of pupil attainment points to the possibility that differences in headteacher practice, which is also likely to differ with age, are also likely to come into play in influencing significant variations in pupil attainment.

Discussion

Similar to children in the UK, Maltese pupils are also able to attain mathematically (Askew et al., 1997; Mujis & Reynolds, 2000; Sammons et al., 2004; Sylva et al., 1999) if conditions at school are favourable for learning (Duncan et al., 2007). However, due to variations in the quality of educational conditions, not all pupils are able to achieve similarly (Mortimore et al., 1988). MIMPS discovered that the differential effect of education with regards to pupil attainment at age 6 was mainly found to be associated with differences in the quality of learning support at the level of the pupil coupled with the quality and frequency of teachers' instructional behaviours at the level of the classroom. The possibility cannot be ruled out that *head teacher age* is a stand-in variable for other broader and qualitatively diverse factors connected with the role of headship in contrast to that of leadership (Bush, 2003; Leithwood, 2003). Locally, classroom and school-based factors, particularly those instructional, also come together in ways that compensate, albeit in part, for the disadvantaging influence usually associated with differences in pupil background factors such as socio-economic status. This implies that pupil background factors are an important but not sufficient condition (Berliner, 1985) in guaranteeing increased or decreased levels of pupil attainment. In spite of differential effects associated with pupil background, the importance of instruction and schooling are highlighted by the influence of instructional factors that exert a compensatory effect.

The Key Role of the Headteacher

Headteachers are known to be important agents in conditioning a positive organisational and academic environment at school conducive for effective teaching (Cotton, 2002; Creemers & Kyriakides, 2008; Levine & Lezotte, 1990; Marzano, 2000, 2003; Mortimore et al., 1988; Sammons, 1999; Scheerens & Bosker, 1997). The finding that younger headteachers exert a more positive effect for pupil attainment than their older counterparts refers, albeit indirectly, to the key role of the headteacher. Headteachers are central in conditioning school-wide leadership, shared and sustained vision, positive and collegial relationships alongside with quality practice (Sammons, 2006). Though unusual for headteacher age to predict pupil attainment, it is not unusual for younger head teachers to act more autonomously, a key feature of leadership (Earley et al., 2012). Earley and his colleagues (ibid.) argue that head teachers have a 'shelf life' and more likely to become disenchanted the longer they are in service.

The Central Role of Teacher Instruction

The findings of the current study show that similar to what was elicited by Marzano (2003) and by Mujis & Reynolds (2000) curriculum coverage and teacher behaviours are influential in promoting pupil achievement. Increased curriculum coverage, coupled with teachers' display of pupils' work, teacher ability to limit disruptions, to establish a positive classroom environment and to manage a positive reward system are all significant predictors of pupil attainment. Time spent delivering lesssons of mathematics was not elicited as a significant predictor of pupil attainment. This is unlike what has been elicited by Campbell et al. (2004) and by Mujis & Reynolds (2001) with older pupils in primary school. Yet, in line with the findings of Sammons et al. (2002) who argue about the importance of quality adult-child interaction for younger pupils. This implies that locally the influence of 'time' is likely to be diffused by quality teacher-pupil interaction; at least during the earlier stages of primary education. However, this does not necessarily imply that time is not indirectly influential. In the classroom, the quality of interaction is known to be latently conditioned by the quality and frequency of teacher behaviours and the amount of time dedicated to lessons during the various lesson stages (Mujis & Reynolds, 2001).

Monitoring Pupil Outcome Over Time

It is through educational policy that issues of quality and equity are attended to and "attempts to define equality and equity in education draw on notions of social justice and social inclusion" (Sammons, 2006, p.3). Gillborn & Youdell (2000) argue that the tenets of social justice refer to: equality of access

to educational provision, equality of circumstance of educational provision, equality of participation when experiencing education and equality of educational outcome. However, is it possible to define and promote quality education in a more equitable manner, if neither pupil outcome nor classroom and school-based processes are systematically monitored? Educational conditions at the school and at the classroom level and the associated impact and influence for pupil attainment, do not operate within a vacuum. The conditions that effect quality of educational leadership, quality of teaching and pupil achievement are dependent on conditions at the policy level (Kyriakides, Creemers & Antoniou, 2009).

More recent policy developments such as those connected with the 'bench-marking' of pupil attainment and the 'age-banding' of pupils in primary classrooms attest to the importance that is currently assigned to the gauging of quality of education provision rather than to the measuring of the associated impact and influence. To remedy the current policy limitations, the author proposes two developments. First, the introduction of baseline assessment. Second, the ongoing and systematic monitoring of pupil outcome.

Baseline Assessment

Consensus exists that assessment is an important feature of sound educational practice (Wilkinson et al., 1998). Importantly, it serves as an initial personal benchmark against which to judge the future performance of children against their earlier performance (Sammons & Smees, 1997). Baseline and initial modes of pupil assessment should also sustain the development of much-needed indices for the identification of more vulnerable groups of pupils who may be at increased risk of experiencing learning delay at school (Sammons et al., 2002; Sylva et al., 2004). Said (2013) recommends the introduction of a system of baseline assessment for pupils at Year 1 (age 5), in the cognitive (language and number) as well as in the social/affective domain.

Summative and Formative Modes of Assessment

Coupled with a national system of end-of-year standardised testing from Year 2 (age 6) until Year 6 (age 11), regular and standardised records of pupil attainment would allow the construction of value-added, or pupil progress scores, which would complement as well as extend both 'bench-marking' and 'age-banding'. It would provide schools with information that facilitates more detailed examination as to the pattern of pupil achievement; for which educational professionals are responsible for. Whilst extremely useful in examining trends relating to pupil outcomes, summative assessment alone cannot yied richer accounts regarding the quality of educational, organisation

and instructional processes associated with both pupil attainment and pupil progress. Consequently, the author also recommends the introduction of a complementary system of formative assessment that includes keeping school and classroom-based records that reflect how pupils have learnt and how they have achieved their targets for learning. If conducted in age-appropriate ways, the systematic implementation of formative forms of assessment should clarify the connection between implicit and explicit forms of teacher knowledge (Nonaka & Takeuchi, 1995) and focus teacher and instructional attention on shorter, daily and more informal cycles of pupil assessment. The kind of assessment that is likely to impact most on pupil achievement (Wiliam, 2009). The kind of assessment that is more likely to be affiliated with effective teacher practice in the classroom (Ko & Sammons, 2010; Sammons, Day & Ko, 2011). The kind of practice that is more likely to occur within school and classroom communities in which pupils learn effectively (Watkins et al., 2007).

Way Forward

In spite of noticeable differences in pupil achievement and in the quality of schooling, teaching and learning in Maltese primary schools, the effects of education remain, as yet, unexamined in terms of larger-in-scale, longitudinal research that quantifies and qualifies the change in direction in pupil achievement, in tandem with the change in direction in the quality of educational provision. In MIMPS, the pupil in classrooms in school model presented in this paper has been applied in a more limited fashion than what is being recommended. Yet, the educational effectiveness template upon which it is based may serve for a variety of uses. One of the possible applications of the pupils in classrooms in schools model within the theoretical framework of Creemers (1994) would be to establish a structure for the systematic and ongoing monitoring of the contexts and processes, associated with repeated measures of pupil achievement (attainment and progress) and teacher, school and policy effectiveness.

References

Askew, M., Brown, M., Rhodes, V., Johnson, D. & Wiliam, D. (1997). *Effective teachers of numeracy*. London: King's College for the Teacher Training Agency.

Anders, Y., Sammons, P., Taggart, B., Sylva, K., Melhuish, E. & Siraj-Blatchford, I. (2010). The influence of child, family, home factors and preschool education on the identification of special educational needs at age 10. *British Educational Research Journal, 1:*1-21.

Berliner, D. C. (1985). The necessary but not sufficient conditions for developing exemplary schools. In Austin, G. R., & Garber, H. *Research on exemplary schools.* Orlando, Florida: Academic Press**:** 127-154.

Brannen, J. (2005). Mixing methods: The entry of qualitative and quantitative approaches into the research. *International Journal of Social Research Methodology, 8(3)*, 173-184.

Bush, T. (2003). *Theories of educational management*. London: Sage.

Byrne, B. (2001). *Structural equation modeling with AMOS*. New Jersey/London: Lawrence Erlbaum Associates.

Camilleri, L., Cefai, C. & Cooper, P. (2011). Social, emotional and behavioural difficulties in Maltese Schools: A multilevel model. *Journal of Research in Educational Sciences*. *Volume 11, 1*(3), 3-15.

Campbell, J., Kyriakides, L., Mujis, D. & Robinson, W. (2004). *Assessing teacher effectiveness: Developing a differentiated model.* London and New York: Routledge Falmer/Taylor & Francis Group.

Carroll, S. J. (1963). A model of school learning. *Teachers' College Record, 64,* 723-733.

Cohen, J. (1988). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20:* 37-46.

Coleman, J. S., Campbell, E., Hobson, C., McPartland, J., Mood, A., Weinfield, F., & York, R. (1966). *Equality of educational opportunity*. Washington D C, U S Government Printing Office.

Comrey, A. L. & Lee, H. B. (1992). *A first course in factor analysis*. Hillsdale, NJ: Erlbaum.

Creemers, B. P. M. (1994). *The effective classroom*. London: Cassell.

Creswell, J. W. (2009). *Research design: Qualitative, quantitative and mixed methods approaches*. Thousand Oaks California: Sage.

Darling-Hammond, L., Cook, C., Jaquith, A. & Hamilton, M. (2012). *Creating a comprehensive system for evaluating and supporting effective teaching*. Stanford, California: Stanford Center for Opportunity Policy in Education (SCOPE).

Day, C., Sammons, P., Stobart, G., Kington, A. and Qing, G. (2007). Teachers matter: Connecting lives, work and effectiveness. McGraw Hill: Open University Press.

Duckworth, K. (2007). *What role for the three R's? Progress and attainment during primary school*. Centre for Research on the Wider Benefits of Learning.

Duncan, G.J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A.C., Klebanov, P., Pagani, L., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K. & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43*, 1428-1446.

Earley, P, Higham, R, Allen, R, Allen, T, Howson, J, Nelson, R, Rawar, S, Lynch, S, Morton, L, Mehta, P and Sims, D. (2012). *Review of the school leadership landscape*. Nottingham: National College for School Leadership.

Gierl, M. J., Rogers, W. T. & Klinger, D. (1999). *Using statistical and judgemental reviews to identify and interpret translation DIF.* Paper Presented at the Annual Meeting of the National Council on Measurement in Education (NCME) at the Symposium entitled, "Translation DIF: Advances and application", Montreal, Quebec: Canada.

Gillborn, D. & Youdell, D. (2000). *Rationing education: Policy, practice, reform and equity*. Open University Press/McGraw Hill Education.

Goldstein, H. (1987). *Multilevel models in educational and social research*. London: Charles, Griffin & Company/New York: Oxford University Press.

Goldstein, H. (1998). Models for reality: New approaches to the understanding of educational processes. *Professorial Lecture Institute of Education University of London*.

Goldstein, H. & Spiegelhalter, D. (1996). League tables and their limitations: Statistical issues in comparisons of institutional performance. *Journal of the Royal Statistical Society, 159,* 385-443.

Gorard, S. & Taylor, C. (2004). *Combining methods in educational research*. London: Open University Press.

Gorsuch, R. L. (1983). *Factor analysis*. Hillsdale, NJ: Lawrence Erlbaum.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Abington: Routledge.

Jencks, C. S., Smith, M., Ackland, H., Bane, M. J., Cohen, D., Gintis, H., Heyns, B. & Nicholson, S. (1972). *Inequality: A reassessment of the effect of family and schooling in America.* New York: Basic Books.

Kim, M. (2001). Detecting DIF across the language groups in a speaking test. *Language Testing, 18*(1), 89-114.

Ko, J. & Sammons, P. (2010). *Effective teaching: A review of research*. University of Oxford, Department of Education.

Kreft, I. (1996). *Are multilevel techniques necessary? An overview including simulation studies*. Los Angeles: California State University.

Kyriacou, C. (1997). *Effective teaching in schools*. Basil Blackwell.

Kyriakides, L., Campbell, R. J. & Gagatsis, A. (2000). The significance of the classroom effect in primary schools: An application of Creemers' Comprehensive Model of Educational Effectiveness. *School Effectiveness and School Improvement, 11*(4), 501-529.

Kyriakides, L., Creemers, B. & Antoniou, P. (2009). Teacher behaviour and student outcomes: Suggestions for research on teacher training and professional development. *Teaching and Teacher Education*, 25, 12–23.

Leithwood, K. (2003). *What we know about successful school leadership*. Prepared for the Task Force of Developing Leadership (Division Administration of the American Educational Research Association.

Maas, C. J. M. & J. Hox (2001). *Sample size for multilevel modelling*. Utrecht, Netherlands: Utrecht University.

Marzano, R. J. (2003*). What works in schools: Translating research into action*. Alexandria, VA: Association for Supervision and Curriculum Development.

Mifsud, C., Milton, J., Brooks, G., & Hutchison, D. (2000). *Literacy in Malta*. London: NFER.

Mifsud, C., Grech, R., Hutchison, D., Morrison, J., Rudd P. & Hanson J. (2004). *Literacy for school improvement: Value-added for Malta*. Agenda.

Mifsud, C., Richard, S. & Hutchison, D. (2005). *Mathematics in Malta: The national survey of Year 1 pupils*. Malta, Education Division, Department of Planning and Development.

Ministry of Education and Employment. (December, 2012). *A National Curriculum Framework for All 2012*. Salesian Press, Malta.

Mortimore, P., Sammons, P., Stoll, L., Lewis, D. & Ecob, R. (1988). *School matters: The junior years*. Open Books

Muijs, D. & Reynolds, D. (2000). School effectiveness and teacher effectiveness in Mathematics: Some preliminary findings from the evaluation of the Mathematics Enhancement Programme (Primary). *School Effectiveness and School Improvement, 11*(3), 273-303.

Muijs, D. & Reynolds, D. (2001). *Effective teaching: Evidence and practice*. Paul Chapman Publishing.

Mullis, I.V.S, Martin, M. O., Foy, F. in collaboration with Olson, J.F., Preuschoff, C., Arora, A. and Galia, J. (2008). *Trends in International Mathematics and Science Study*. Lynch School of Education, Boston College: TIMSS & PIRLS Study Centre.

Mullis, I.V.S., Martin, M. O., Foy, P. & Arora, A. (2012). *Trends in International Mathematics and Science Study*. Lynch School of Education, Boston College: TIMSS & PIRLS Study Centre.

Mullis, I.V.S., Martin, M.O., Kennedy, A.M., Trong, K.L., & Sainsbury, M. (2009). *PIRLS 2011 assessment framework*. Chestnut Hill, MA: Boston College.

National Statistics Office. (2005). Census 2005. Valletta: Malta, National Statistics Office.

Neely, A., Gregory, M. & Platts, K. (1995). Performance measurement system design: literature review and research agenda. *International Journal of Operations & Production Management, 15*(4), 80-116.

Nonaka, I. & Takeuchi, H. (1995). *The knowledge-creating company.* New York: Oxford University Press.

Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical linear models: applications and data analysis methods*. London: Sage.

Reynolds, D., Creemers, B., Stringfield, S. & Schaffer, G. (2002). Creating world-class schools: What have we learnt? In Reynolds, D., Creemers, B., Stringfield, S., Teddlie, C. & Schaffer, G. (2002). *World Class Schools. International Perspectives on School Effectiveness*. London/New York: Routledge and Falmer Press.

Sammons, P. (2006). *Embracing diversity: New challenges for school improvement in a global learning society*. Keynote speech. Fort Lauderdale, Florida, 4th January 2006: International Congress for School Effectiveness and School Improvement.

Sammons, P., Day, C. & Ko, J. (2011). Exploring the impact of leadership on pupil outcomes: Results from a study of academically improved and effective schools in England. *International Journal of Educational Management, 25*(1), 83-101.

Sammons, P. & Smees, R. (1997). *Developing value-added models at KS1 using baseline screening at entry to school*. London: Institute of Education, University of London.

Sammons, P., Smees, R., Taggart, B., Sylva, K., Melhuish, E., Siraj-Blatchford, I. & Elliot, K. (2002). *EYTSEN Technical Paper 1: Special educational needs across the pre-school period*. London: Institute of Education.

Said, L. (2006). Influences on pupils' mathematical outcome: a description of challenges. *Educate, 6*(2), 42-60.

Said, L. (2013). *An examination of the pupil, classroom and school characteristics influencing the progress outcomes of young Maltese pupils for mathematics*. Unpublished Doctoral Thesis. London: Institute of Education

Scheerens, J. (1992). *Effective schooling: Research, theory and practice*. London: Cassell.

Seidel, T. & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: the role of theory and research design in disentangling meta-analysis results. *Review of Educational Research, 77*, 454.

Sireci, S. G. (1997). Problems and issues in linking assessments across languages. *Educational Measurement: Issues and Practice, 16,* 12-19.

Schagen, I. (1990). A method for the age standardisation of test scores. *Applied Psychological Measurement, 14,* 387-393.

Snijders, T. A. B. (2005). Power and sample size in multilevel linear models. In Everitt, B. S. & Howell, D. C (eds.). *Encyclopedia of Statistics in Behavioural Science*, *3*, 1570 – 1573, Chicester:Wiley.

Snijders, T. A. B. & Bosker, R. J. (1999). *Multilevel analysis: an introduction to basic and advanced multilevel modelling*. Thousand Oaks: New Delhi, London.

Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I. & Taggart, B. (2004). *Technical Paper 12 The Final Report: Effective pre-school education*. London: Institute of Education.

Sylva, K., Sammons, P., Melhuish, E., Siraj-Blatchford, I. & Taggart, B. (1999). *The effective provision of preschool education (EPPE) project*. London: Institute of Education.

Tashakkori, A. & Creswell, J. (2007). The new era of mixed methods. *Journal of Mixed Methods Research, 1*(1), 3-7.

Tashakkori, A. & Teddlie, C. (2003). *Handbook of mixed methods in social & behavioural research*. Thousand Oaks: Sage.

Teddlie, C., & Tashakkori, A. (2003). Major issues and controversies in the use of mixed methods in the social and behavioural sciences. In A. Tashakkori, & C. Teddlie (Eds.), *Handbook of mixed methods in social & behavioral research*. Thousand Oaks, CA: Sage.

Teddlie, C. & S. Stringfield (1993). *Schools make a difference*. New York: Teachers College Press.

Tymms, P., Merrell, C. & Henderson, B. (1997). The first year at school: a quantitative investigation of the attainment & progress of pupils. *Educational Research and Evaluation, 3,* 101-118.

Watkins, C., Carnell, E. & Lodge, C. (2007). *Effective learning in classrooms*. London: Paul Chapman Publishing.

Wiliam, D. (2009). *Assessment for learning: why, what and how?* London: Institute of Education.

Wilkinson, J. E., Watt, J., Napuk, A. & Normand, B. (1998). *Baseline assessment literature review and record-keeping in Scotland*. The Scottish Office and Education Industry Department (SOEID): Educational Research Unit.

Yamane, T. (1967). *Statistics, an introductory analysis*. New York: Harper and Row.

Zumbo, B. D. (1999). *Logistic regression modelling as a unitary framework for binary and Likert-type (ordinal) item scores: A handbook on the theory and methods of differential item functioning (DIF)*. Ottawa, Canada: Directorate of Human Resources Research & Evaluation National Defense Headquarters.