A large-scale quantitative investigation of teacher-feedback and students' on-task behaviour as associated indicators of the social-emotional climate for learning in academic lessons in UK secondary schools using a systematic observation method: 'MICRO'

Brian John Bishop Apter

DEdPsy Thesis

SN: 0842733

16th September 2016

School of Psychology

University of Cardiff

Dedicated to the memory of my dad, John Henry Apter, who told me: "If a job is worth doing, it is worth doing properly."

Acknowledgements

This thesis would not have been possible without the help and encouragement of some important colleagues who are also very good friends: Dr Jeremy Swinson, Dr Francesco Sulla, Ian Jones, Helen Davies, Bill Bennett and Claire Gilbert. I am indebted to my daughter Dr Nora Apter for her encouragement, support, intelligence and wise counsel; and to my daughter Harriet for challenging me, forgiving me and making me laugh. Neither would this thesis have been possible without the encouragement, support, occasional impatience, frequent tolerance and love of my partner and wife, Lisa Apter, and also of my mother, Iris Apter. I must also mention with huge gratitude the gentle scholar, gentle man and UFO spotter, Dr Simon Griffey who was my commissioning tutor at the University of Cardiff and an encouraging critical friend in the early days of this project - I wish very much he was here to baptise the baby. Thanks therefore to Dr Ian Smillie and colleagues at the University of Cardiff for stepping-up in Simon's absence.

A large-scale quantitative investigation of teacher-feedback and students' on-task behaviour as associated indicators of the social-emotional climate for learning in academic lessons in UK secondary schools using a systematic observation method: 'MICRO'

Brian Apter[1] – University of Cardiff

SUMMARY

Systematic observations by 33 psychologist-observers of 228 lessons in 28 UK secondary schools were included in this study. A new method: the Mixed Interval Classroom Observation (MICRO) schedule was used to collect data about teachers' use of 4 different types of verbal feedback and make associations with the proportion of students' time in academic classes that they were following teachers' directions and were 'on-task'.

Key findings from 27 UK secondary schools included: students were significantly less 'on-task' than students in UK primary schools; secondary school teachers used low frequencies of positive verbal feedback directed towards academic work and behaviour and much higher frequencies of critical comments directed towards behaviour; teachers' critical comments directed towards behaviour were significantly associated with lessons where students were less compliant with teachers' directions; and teachers who used high frequencies of positive comments directed towards academic work and social behaviour were not associated with lessons where students followed teachers' directions more. The number of teachers who did not use any positive comments about social behaviour was high compared to the findings of previous researchers. Teachers who used verbal feedback were more likely to use more with the lower year groups. Unlike primary students, no evidence was found that secondary students were more engaged with academic work when taught by teachers who used higher levels of verbal teaching behaviour: teachers who talked more.

A number of contextual factors were also examined for their association with students' compliance with teachers' directions. Findings included: teachers who were more experienced were more likely to be teaching students who followed their directions.

Subsequently, every teacher in one particular secondary school that had been placed in 'special measures' following an Ofsted inspection, was observed twice, A.M. and P.M., using the MICRO schedule, and the results were statistically compared with the UK dataset of 27 secondary schools described above. Initial findings of this exercise were shared with the school's senior leadership team (SLT) and their discussion was recorded, transcribed and analysed using the 'Iterative Learning Conversations' (ILCS) discourse analysis method (Apter, 2014). Findings revealed that the SLT believed that the exercise of exploring the statistical comparison was of significant utility in strategic planning. The results provided evidence for their beliefs as to why the maths department had been found to have serious weaknesses during the inspection, and that the way that teachers used verbal feedback throughout the school required further monitoring and improvement.

Conclusions are drawn about the nature of teachers' verbal feedback in secondary schools and how Bandura's (1977) Social Learning Theory and reciprocal determinism provide a robust psychologically explanatory model (more-so than a behavioural stimulus-response model) as to how the bi-directional interaction of teacher-feedback and student behaviour works to indicate the conduciveness of the social-emotional climate for learning in a school.

[1] Email: apterchapter@gmx.com

CONTENTS **Page**

[2] In the research literature, the terms: *systematic observation*, *structured observation* and *interaction analysis* are not functionally differentiated (O'Leary, 2014, p.50). 'Systematic observation' has been used here as it is the most frequently used of the 3 terms in the cited literature.

CONTENTS (continued) **Page**

TABLES AND FIGURES

TABLES AND FIGURES (continued)

Chapter 1 - Introduction

- **The role and attributed importance of teachers' verbal feedback in modern secondary school classrooms**
- **The importance of systematic classroom observation in the applied psychology of Educational Psychologists**

1.1 The role and attributed importance of teachers' verbal feedback in modern secondary school classrooms

Whilst Boyle (2009) noted: "The great thing about primary education is the positivity and praise that kids get" (p.33), teacher-feedback is attributed to be a significant component of teaching and learning in *both* primary and secondary schools. Hattie (2009) does not differentiate between primary and secondary when reporting the effect sizes of a range of interventions from more than 800 meta-analyses - teacher-feedback was 10th in a league table of 138 intervention effect sizes (d= 0.73)[3]. Extrapolating a league table of exclusively *teacher* effects, feedback was 3rd. 'No. 1' was 'providing formative evaluation' (d= 0.90), which might also be classed as feedback by a classroom-observer; and no. 2: 'micro teaching' (d= 0.88), which is where lessons are videoed to provide constructive feedback to teachers.

Curiously, Hattie qualified his judgement of teacher-feedback, thus: "Programmed instruction, praise, punishment and intrinsic rewards were the least effective forms of feedback for enhancing achievement." (p.174)

This marked a departure from much of the received behavioural wisdom as it had been researched and applied in schools over the last 50 years or more that 'praising' desired learning-behaviour would increase the frequency with which it occurs and thus improve learning-outcomes, providing that the praising-comment was constructed with specificity: explicitly describing the desired-behaviour and identifying the student by name who was using the desired-behaviour (*Assertive Discipline*, Canter and Canter, 1976).

1.2 The *Sutton Trust* constructed a meta-analysis: the *Teaching and Learning Toolkit* (Higgins, Katsipataki, Kokotsaki, Coleman, Major, and Coe, 2014) which identified which teaching interventions worked best. Comprised of comparisons of intervention-effectiveness

[3] *'An effect size of d= 1.0 indicates an increase of one standard deviation on the outcome – in this case the outcome is improving school achievement. A one standard deviation increase is typically associated with advancing children's achievement by two to three years*, [or] *improving the rate of learning by 50%'* (Hattie, 2009, p. 7)

and cost-benefit analysis, impact measurement was tabulated in terms of how many additional months of progress an average student would make if a specified intervention were invested in by a school. Feedback to students by teachers was reported to be an inexpensive investment for a school and an effective way of positively enhancing educational outcomes with up to *'plus-8 months'* achievement gain – the maximum gain registered in the league table.

1.3 It is also worth noting Higgins *et al.'s* (2014) qualification for what constitutes the most effective teacher feedback according to the *Sutton Trust:*

> Providing effective feedback is challenging. Research suggests that it should be specific, accurate and clear (e.g. "It was good because you..." rather than just "correct"); compare what a learner is doing right now with what they have done wrong before (e.g. "I can see you were focused on improving X as it is much better than last time's Y…"); encourage and support further effort and be given sparingly so that it is meaningful; provide specific guidance on how to improve and not just tell students when they are wrong; and be supported with effective professional development for teachers. (*Sutton Trust* website)

1.4 Neither Hattie (2009) nor Higgins *et al.* (2014) differentiated between how verbal feedback might be received in different ways by different age-groups. The likely-responses of modern 6, 11, and 16 year-olds in the UK to being told that they are 'working hard' would be quite different. An 11 year-old or a 16 year-old might consider this patronising, whereas a 6 year-old would probably enjoy such attention. The author's previous findings (Apter, Arnold and Swinson), 2010) agreed with previous UK researchers (Merrett and Wheldall, 1987; Harrop and Swinson, 2000) that positive feedback directed at academically-oriented behaviours was an effective device with primary-phase students and that it was associated with increased 'on-task' time in lessons.[4] The current research comes to different conclusions in respect of how feedback works with secondary-phase students (Chapters 4 and 5).

---

[4] Hattie also uses the term: *on-task* to refer to 'practising' and 'practice' rather than the usage employed by the current research and previous research (Wheldall, *et al.,* 1985; Apter, *et al.,* 2010) where 'on-task' means: the student is observed to be behaving as if they are complying with the teacher's last instruction - which might have been: 'Sit still, and listen quietly.'

1.5 Carol Dweck (1986, 1999), a North American psychologist who has been concerned with how North American parents, teachers, lecturers and sports-coaches use 'praise' to tell children and young people that they are 'clever' or 'physically able'; rather than 'industrious' and 'persistent', has written voluminously on this subject. Dweck's conclusion is that positive encouragement directed at effort and particularly persistence works much better and is: '*The Secret to Raising Smart Kids*' (Dweck, 2015).

There is a cultural-linguistic specificity to the content of 'praise' statements and their effectiveness in North American educational institutions that is outside the scope of the current UK research presented here. Whilst not wishing to ignore within-person personality theories, this research and its literature review cannot do justice to a proper examination of motivational research, or indeed, incorporate popular metacognitive theories about what characterises the most effective 'praise' statements in North America. Nonetheless, Dweck's research is briefly returned to in the discussion of findings (Chapter 6).

1.6 This research examines predictive associations between the time that students spend following directions that teachers give them in academic lessons in UK secondary schools (the 'On-task%' dependent variable) and different types of teacher-feedback they are given using systematic quantitative classroom observation.

The current research also attempts to examine a number of correlations, for example: between following teacher' directions; the teacher's gender; and their years of teaching experience by quantitatively exploring these associations using the observations of a large sample of lessons observed from across the UK by psychologists; and by using a second comprehensive sample whereby every teacher in a single UK secondary school was observed teaching and giving feedback to students twice – once in the morning and once in the afternoon. The two datasets were then used as comparators with each other in order to enrich and support or criticise and defeat a range of hypotheses being tested; and assist in answering research questions posed.

1.7 The importance of systematic classroom observation in the applied psychology of Educational Psychologists

The psychological systematic observation of classrooms to examine and enhance the way in which teachers teach and learners learn in school lessons is arguably one of two original, persistent and pervasive themes that ensured the timely birth in Europe and North America

of the domain of applied psychology that is called 'educational psychology'[5] (Thorndike, 1903) at the end of the nineteenth century (Chapter 2). The other was the testing and measurement of children's mental abilities (Thorndike, 1903; Binet and Simon, 1916; Terman, 1916; Burt, 1917) in order that specialised or segregated provision could be decided upon for the 'Educationally Sub-Normal - Mild or Severe' (Warnock 1978) or in order that learners could be designated - preposterously by today's standards – 'ineducable' (Segal, 1968).

It is not intended to imply that school-based educational psychology that uses systematic classroom observation is the most significant development of educational psychology, but whilst it is a limitation upon scope imposed by the design of the current research and the research standpoint, it is possibly not a limitation on its more general relevance to applied psychology, and to educational research; for example: as it applies to child and adolescent development; to child and adolescent behaviour; and to learning and behaviour in school.

1.8 Various ways in which Educational Psychologists[6] and other educational researchers have tried to develop valid and reliable methods for systematically observing classrooms for the purpose of examining and evaluating teaching and learning are described in Chapter 2. This review of the research literature serves to justify and predicate the current research. It also serves to show the linkage between this type of research and behaviourism (Watson, 1913) but also its disconnection; and its connection instead with a social-cognitive consideration of the social emotional climate for learning in school-classrooms (Bandura, 1986). Chapter 3 describes epistemological and methodological considerations of the current research and the quantitatively-lead method that has been used. Chapter 4 and Chapter 5 are reports of the results of the current research. Chapter 6 is a discussion of findings in respect of the research questions and hypotheses being examined with reference to current education and educational psychology contexts in the UK. Chapter 7 is a brief set of conclusions and future directions.

1.9 What is new in this study? The range and depth of the themed research-literature review (Chapter 2) of systematic classroom observation is more comprehensive than has been attempted previously in order to contextualise and predicate the current research. The research uses a new method, the *MICRO* observational tool (Chapter 3, Appendix) – a

[5] Thorndike coined the term 'Educational Psychology' as a book title in 1903.
[6] The UK's Health Care and Professionals Council (HCPC) list 'Educational Psychologist' as a protected title. For this reason it is capitalised where it occurs.

development of a number of previous different time-sampling approaches cited in the literature (Chapter 2) – to assist observers in collecting data from their observations of secondary schools lessons. The research examines a larger sample of 28 UK secondary schools than has previously been schieved with approximately 238 lessons being observed in total. The range of statistical measures used to examine and compare the 2 datasets used in the study: UK27 - observational data about 106 academic lessons from 27 different UK secondary schools; and UK1 – observational data about 122 academic lessons from one UK secondary school, was more robust and critically appraised than has been previously attempted (Chapters 4, 5 and 6). The quantitative findings of the research were critically examined using a new qualitative approach – Foucauldian Iterative Learning Conversations (Apter, 2014; Chapters 3 and 5) with the senior leadership team of the UK1 secondary school.

1.10 In terms of introducing this research it is important to describe the author's professional and theoretical standpoint. The author qualified as a teacher in 1979 and specialised in teaching and working with young people who had committed serious criminal offences. Since 1996, the author has been a professionally qualified Educational Psychologist who for many years worked in UK local authority schools using a range of psychological paradigms and approaches in his work as an applied psychologist and specialist in challenging behaviour, including: psychotherapeutic (Winnicott, 1976) ; human-centred (Rogers , 1951) and critical-psychological (Foucault, 1969) approaches; but the dominant and recurring psychological academic and professional theme has been Social Cognitive and Social Learning Theory; and social-cognitive and behavioural applied psychology (Bandura, 1977; 1986).

Chapter 2 – Review of research literature, research questions and hypotheses

- **Five stranded historical thematic review of systematic[7] classroom observation**
- **The early beginnings of educational psychology: Thorndike at the Teachers' College, Columbia University, and a psychological approach to studying the interaction between teaching and learning**
- **The North American psychological-empiricist strand: experiments attempting to measure the effect of verbal praise and reproof used by teachers in classrooms**
- **The transatlantic pedagogic strand: real-time observation methods of teaching and learning using coding tables to tally classroom interactions in order to promote the social-emotional climate of classrooms and appraise teachers**
- **The North American (and New Zealand) observational (non-empirical) educational psychology strand**
- **The UK observational (non-empirical) educational psychology strand: real-time 'naturalistic' time-sampling methods to record classroom interactions**
- **Research questions and hypotheses**

2.01 Five stranded historical thematic review of systematic classroom observation

Presented here is a 5-stranded review of psychologically-informed classroom observation; particularly: previous research that has used systematic quantitative observations of whole classrooms to examine naturally occurring verbal behaviour of teachers, their use of approval and disapproval, and associations made by researchers between this and the learning behaviour of school students.

This is not an exhaustive systematic review. In terms of search strategy and what has been included and what has been left out, a archaeological-narrative approach (Foucault, 1969) has been used to produce a *descriptive* discourse (Khoo, Na and Jaidka, 2011). The aim is to involve the reader in key developments by providing a historical-psychological context of research that is discussed. Unlike an *integrative* review (ibid.), more detail is provided about the process of each study rather than an exclusive focus upon results. The process has been selective and has focussed primarily on British and North American research and its history[8].

[7] In the research literature, the terms: *systematic observation*, *structured observation* and *interaction analysis* are not functionally differentiated (O'Leary, 2014, p.50). 'Systematic observation' has been used here as it is the most frequently used of the 3 terms in the cited literature.

[8] A number of search engines have been used including: ERIC, PsycINFO, PsycArticles and Google Scholar. Original sources have usually been procured in the case of books published prior to 1960.

2.02 The early beginnings of educational psychology: Thorndike at the Teachers' College, Columbia University, and a psychological approach to studying the interaction between teaching and learning

2.03 The history of recording and using systematic classroom observations to measure the effect of different types of verbal behaviour used by teachers and students, and the effect that teacher verbal behaviour has had on student learning, extends back more than 100 years (Wragg, 1999). It is probable that a less systematic scrutiny of teachers' verbal behaviour and its association with students' learning behaviour goes much further back to the founding of the first schools (e.g. Chengdu Shishi school, China 143-141 B.C.; and in Europe, the King's School, Canterbury, 597 A.D.); and universities in the 11th century (e.g. University of Bologna, Italy, 1088) as the relationship between teaching and learning began to be examined when teachers were trained (Jones, 1924; Monroe, 1928). Earlier still, Plato and Xenophon (c. 400-350 BCE) describe Socrates verbal behaviour as a teacher with much admiration and analysis. In their written descriptions of Socrates' philosophical practice – typically the *Apology* (c. BCE 399-390) by Plato and *Memorabilia* by Zenophon (c. BCE 371-365) - they depict Socrates and the 'Socratic' method of asking questions to elicit a train of enquiry and learning. These accounts were arguably the first examples of observations of a teacher and teacher-talk, albeit that the classroom was an Athenian forum or street corner.

2.04 From 1840 onwards, in England, Scotland and Wales, largely unregulated schools intended to educate the children from paupers' homes to read well-enough to manage the scriptures and perform simple arithmetic multiplied (Sturt, 1967). Teachers were trained through unregulated supervision and apprenticeship (Painter, 1896; Jones, 1924; Rich, 1933), the well-meaning agency of Teacher Pupil Centres, and the homilies of National School Society's schools' inspectors.

There was little evidence of the observation of teaching being used for a training purpose until the establishment of one of the earliest teacher training colleges in Scotland: the Normal School, conceived and developed by David Stow in 1836 under the auspices of the Glasgow Educational Society. Scripted model lessons were enacted and observed by trainees who were indoctrinated with the principles of Stow's teaching method (Sturt, 1967). Stow's method and many training concepts were subsequently imported from the Normal School into England by Sir James Kay-Shuttleworth - notably in the establishment of St Mark's College, Chelsea, in 1841. It was a prescribed element of training that a trainee

teacher: "Observes, teaches under observation, writes a journal and record of observations, which must be submitted weekly to the rector for his perusal and criticism." (p.134)

How observations were conducted and what the desirable features of model lessons were is unclear. Whilst observations informed teacher-training in the colleges and training schools, there is no evidence that teachers, or their trainers, were trained in observation skills.

2.05 The potential benefits of a teacher's use of verbal praise or approval and reproof or disapproval appears to have been unconsidered as a psychological variable until Binet and Vaschide's (1897) experimental conjecture that encouragement might be improving physiological performance by improving 'mental' performance.

In one of a series of experiments in France using primary-age boys as subjects and a device: a *dynamometer*, for measuring their physical effort when performing a simple task with each hand, Binet and his colleague observed that the school's Director (the equivalent of a Head Teacher) encouraged greater effort from his pupils using a series of exhortations:

> Allez! Nous avons vingt-cinq kilogramme., main droite. Allons hardi! Vingt-et-un, main gauche. Vingt-tois, encore main droite. Allons tu peux faire mieux que ça toi. Vingt-et-un, gauche. C'est cette fois que nous arrivons à vingt-six? Nous avons vingt-huit! Vingt-tois, gauche. Allons donc! Vingt-tois, encore, droite. Allons, allons! Vingt-quatre, gauche. Allons, marche, Vingt-cinq droite; il est arrivé tout la meme. Hardi là, hardi; vingt-et-un gauche.[9] (p.34)

Binet and Vaschide commented that they thought that the school Director was attempting an excitation of each subject's pride by using encouragement, irony and sarcasm. This was probably the first example in the psychological research literature of a psychologist noticing a relationship between a pupil's effort and a teacher's verbal behaviour.

2.06 It is unsurprising that in 1897 the two psychometricians were not more interested however. Watson's (1913) treatise: 'Psychology as the Behaviorist Views it', informed by Pavlov's conditioning experiments (Todes, 2014) had not yet been written. Indeed, the

[9] Translated: Come on! We have 25 kg., right hand. Be bold! 21, left hand. Yet 23, with the right hand. Come on you can do better. 2, left. This is the time we get to 26? We have 28! 2, left hand. Come on! 23, more with the right. Come on, come on! 24 left. Come on, right hand, 25; it happened all the same. Come on there, be bold; 21 left hand.

concept of *reinforcement* was not to be fully developed for another 40 years in the work of Skinner (1938) on the other side of the Atlantic.

2.07 Edward Thorndike was a lecturer at the Teacher's College from 1899 to 1939. His early research into the learning of animals at Columbia University (Thorndike, 1898) was identified by Pavlov (1928) as seminal:

> I must acknowledge that the honour of having made the first steps along this path belongs to E.L.Thorndike. By two or three years his experiments preceded ours and his book [Thorndike 1911] must be considered a classic, both for its bold outlook on an immense task and for the accuracy of its results. (Introduction)

Thorndike distanced himself from Watson's radical behaviourism, writing that consciousness and conscious thought in humans was indisputable (Thorndike, 1911). Thorndike published 'Educational Psychology' (1903), and proposed that:

> The work of education is: 1. To supply the needs of the brain's healthy growth and to remove physiological impediments to it. 2. To provide stimuli to desirable mental variations and to withhold stimuli from the undesirable. 3. To make the outcome of desirable activities pleasurable and to inhibit their opposites by discomfort. (p.79)

Thorndike generalised the *Law of Effect* from animal experiments to children's learning, reporting the results from two small experiments with school-age children (Thorndike, 1927). Thorndike's connectionist law proposed that practice, associated with verbal feedback (a teacher saying "Right!" or "Wrong!") increased the strength of a new connection in the brain, thus ensuring future accuracy of new behaviour. Thorndike's Law of Effect was less 'black box' behaviourist (Watson, 1913) and more a precocious precursor to Kandel's neuroscientific finding of how learning occurs in animals' brains (Kandel and Tauc, 1965).

2.08 The psychometrician, James Cattell (Joncich, 1968), was also a member of staff at Teacher's College, Columbia University until 1917; and until 1930, so was John Dewey, the functionalist psychologist and philosopher (Dewey, 1897). Significantly, the first 'time and motion' psychologist, Lillian Gilbreth studied under Thorndike at Teacher's College between 1900 and 1902. Thorndike's influence is evident in her work with husband Frank promoting manufacturing efficiency (Gilbreth, 1917). The Gilbreth's work predicated the subsequent development of classroom observation being used for teacher appraisal and learning efficiency, but it was Thorndike's applied and experimental psychology in

educational settings and his commitment to his professorial role at the college until his retirement in 1939 that contributed to psychology's influence on classroom research from that time (Joncich, 1968; Brooke, Tsuei, and Malinow, 2003).

2.09 Professor Ramiett Stevens studied at Teacher's College, Columbia University at this time. Stevens is credited in a number of reviews (e.g. Hoetker and Ahlbrand, 1968; Wragg, 1994) as having been the first researcher to have made systematic transcriptions of naturally-occurring classroom verbal behaviour. Stevens quantified the records she made in terms of the number of questions asked by teachers, the type and quality of the questions, the way in which questions were used to provoke learning and the extent to which pupils' responses provided evidence that questions were helpful for their learning.

Stevens' study was distinguishable from other research reported here until the research of the 1960s because it was not experimental. It employed a systematic observational method that attempted to capture the naturally-occurring verbal behaviour of teachers and pupils' responses.

For her doctoral study, Stevens (1912) transcribed verbal behaviours of teachers and pupils from 100 lessons she observed in primary and secondary schools in New York. Stevens (ibid., p.22) noted that the amount spoken by teachers and students was *relatively* 64% and 36% [10].

2.10 It was not Stevens' purpose to determine how much time was spent talking. Stevens counted the number of words used by teachers compared to those used by pupils in 20 of her transcriptions to measure proportionality (Stevens, 1912, table 3, p.22). A re-read of Stevens' transcripts with a stop-watch reveals - albeit inexactly - that pupils worked in silence for long periods of observed lessons. Stevens estimated that the amount of time taken up in lessons by questions and answers was 80%: "I believe it is safe to say that eight-tenths of school time is occupied with questions and answers." (p.6)

There is no evidence provided by Stevens (1912) to support this observation and it seems unlikely that Stevens could vouchsafe accurate measurement. It is also interesting to note that Stevens' transcriptions are devoid of pauses, dysfluencies, repetitions or re-phrasings. This suggests that Stevens corrected or modified her own transcripts – her raw data - before

[10] Misinterpreted by Wragg (1994) as if they were proportions of the observed time-period: *'It was found that teachers talked for about 64 per cent of the time and pupils 36 per cent.'* (p.8, Wragg,)

analysing them. Stevens was interested in the type of questions teachers used and whether question-type was a variable that supported or impinged upon pupils' learning. The degree of transcript veracity Stevens adopted was probably sufficient for this purpose:

> The large number of questions [used by teachers] suggests that we are coming, more and more, to make the classroom the place for displaying knowledge instead of a laboratory for getting and using it… …The large number of questions suggests an almost total absence, in the practice of our class rooms of any psychological principles underlying aims or methods. (p.25-26)

2.11 The role of approval and disapproval in Steven's (1912) transcripts is not commented upon and rarely appeared to be used by teachers. In her transcripts, there are three examples of approval recorded:

> Pupil: They had high ideals; they were true to them and were great.
> Teacher: A better answer still. (p.60)
> ---
> Pupil: Self-confident.
> Teacher: That is just what I wanted – self-confident. (p.84)
> ---
> Teacher: You consider Ralph De Wilten the hero?
> Pupil: Yes.
> Teacher: You think it was the times rather than the man himself?
> Pupil: Yes.
> Teacher: That is perfectly true. (p.91)
> ---

There is one example of disapproval:

> Pupil: …but before that he knew that Ralph De Wilton was the Palmer.
> Teacher: That is not very clear; anyone clear that up! Edward! (p.89)

It might be that teachers in New England in 1912 neither used praise nor disapproval to any significant extent in their classroom discourse. Intuitively, it seems unlikely that the latter has ever been absent from teacher discourse.

2.12 Stevens' research inspired subsequent researchers in North America to use classroom observation as a research tool in the early part of the 20th century to repeatedly examine the structure of teacher-questions and their relationship with pedagogy (Colvin, 1919; Monroe

and Carter, 1923; Barr, 1929; Colvin, 1931; Bagley and MacDonald, 1932; Briggs,1935; Dale and Raths, 1945). Questions and answers were considered the main device of teaching.

2.13 The North American psychological-empiricist strand: experiments attempting to measure the effect of verbal praise and reproof used by teachers in classrooms

2.14 Gilchrist (1916) used an experimental methodology rather than naturalistic observation to investigate the effect of 'praise' and 'reproof' used by teachers. A student on a post-graduate experimental-psychology course at Cornell University, Gilchrist was probably the first psychologist to attempt such an experiment.

Two groups of educational-psychology students were given an English language test - the *Courtia English Test 4B.* On completion students in Group A were told that they had performed badly and were required to re-sit the test to see whether they could do better. Students in Group B were told they had performed well but that were required to re-sit the test to see whether they might improve their scores. Gilchrist summarised the results:

> 1. The group that was praised improved the group score by seventy-nine percent in the second test.
> 2. The group that was reproved made a lower group score in the second test than it had made in the first test.
> 3. In the second group those who had done well in the first test did not do so well in the second test, whereas those who had not done well in the first test improved their scores in the second test.
> 4. The percentage of loss and gain express presumably, the effect of the expressions of praise and reproof.
> 5. A teacher may find in the results a means of testing her [Sic.] own tendency to overestimate or to underestimate the effect of praise and reproof. For the extent to which she finds the results to be greater or less than she expected them to be ought to be some indication of her tendency to misjudge the effect of criticism. (p.874)

Gilchrist assumed transferability of his hypothesis to school situations and younger subjects, as was evident in the title of his report: '*The extent to which praise and reproof affect a pupil's work*'. The generalisability of Gilchrist's findings was limited by the convenience of his subject-pool.

2.15 Subsequently, Kirby (1913) at Teacher's College, Columbia, sought to demonstrate practice effects upon the school-attainments of Grade 3 and 4 children. The acknowledgment at the beginning of Kirby's paper reveals that Thorndike had encouraged Kirby and had suggested the subject of this research.

Kirby was later referenced by Hurlock (1925) and as a result, subsequent investigators such as Johannesson (1967), as having explored the effect of 'encouragement' but there is no evidence of this. Kirby writes that children appeared inspired to do better by comparing current with previous performances as they practised.

2.16 Chapman and Feder (1917) of the Western Reserve University in Cleveland reported an experiment concluding that 'incentives' significantly increased the academic output of children. Their subjects, divided into 2 groups - an experimental condition and a control group - were 36 fifth grade boys and girls. The incentives used were graphs of performance publicly celebrated in the class. The authors reported significantly increased the quantity of written output and improved accuracy (fewer mistakes) compared with the experimental group.

2.17 Elizabeth Hurlock, a psychology lecturer at Columbia, completed research (1924, 1925) into the effectiveness of praise and reproof used by teachers in schools. Hurlock obtained results from her control-group designs that supported Gilchrist's hypothesis *in situ* in school. Hurlock (1925) reported that both praise and reproof worked to improve students' results in the short term. In the longer term praise was more effective in producing increased output and improved accuracy. Hurlock reported gender differences: boys seemed more affected by reproof than girls; and girls more affected by praise than boys. She noted no significant effect-size differences by age.

Hurlock (1931) published a review of psychological research from 1890 to 1930 concerned with the psychology of incentives and concluded that praise, rewards and encouragement appeared to be more effective than reproof, sarcasm and punishment in increasing output and improving behaviour across the age-ranges in schools, colleges, prisons and industry.

2.18 Brenner (1934), working under the auspices of Teachers College, Columbia University, criticised previous experimental research investigating the efficacy of praise and reproof noting methodological faults including: small number of subjects; inadequate specification of variables, e.g. verbal comments; inadequate statistical analysis; and

unwarranted generalisation. Brenner concluded that it had not been proven that praise was more effective than reproof.

Brenner's own research utilised a more sophisticated experimental control-group design with nineteen third grade children as subjects and reported that immediate praise was more effective than delayed praise.

2.19 In 1941, Schmidt published an inclusive review of the literature in respect of praise and reproof. Schmidt found that the 28 experimental designs used in the research cited were inadequate for purpose. Schmidt excepted Brenner.

Schmidt used Brenner's work to develop the experimental design he used. Neither Scmidt nor Brenner questioned whether an experimental control-group design was the best way to research the effectiveness of praise and reproof used in the classroom. At this point in the history of psychology, when the aspiration to establish psychology as a science was ascendant, there appears to have been a reluctance to consider alternatives to experimental designs.

Schmidt's experiment attempted to arrive at a definitive position on the relative merits of praise and reproof by using a matched-groups experimental design and correlation statistics. Schmidt arrived at the conclusion that there was no significant difference in the effect size between blame and reproof in different classroom settings. He concluded that biggest differentials in output and accuracy are caused by different experimenters and their different approaches to working with the subjects to set up the experiment.

2.20 In a later review, Kennedy and Willcut (1964) considered 33 research reports from between 1897 and 1964. They reported that by 1940 the dominant conclusion when experimental designs were employed was that praise and blame (reproof) were equally effective as incentives.

Kennedy and Willcut note that there was less research interest in the field between 1940 and 1950 because of the 2nd World War. They identified only 3 reports of any methodological substance and no new revelations. Reviewing 7 experimental reports published between 1950 and 1960, they concluded that there was no new evidence to contradict the established view at that time that there was no difference between the effects of praise and blame upon school age children.

Kennedy dominated an albeit sparse and dwindling experimental field at that time in respect of research about praise and blame (Kennedy, Turner and Lindner.,1962; Tiber and Kennedy, 1964; Kennedy and Willcut, 1964; Kennedy and Vega, 1965; Kennedy and Willcut, 1965). Kennedy's research group concentrated on subsets of the school population such as: particular age-bands; the race of teachers; IQ differences; 'mental defectives'; and blind children, and their respective performances on mechanical tests e.g. sorting tasks when praise or blame was used as an incentive. Apart from one experiment that used an IQ test as the task (Tiber and Kennedy, 1964), praise was consistently identified as a more effective incentive than blame.

2.21 Johannesson (1967), reviewing work in the area including Kennedy's, concluded that there was substantial evidence that:
- Repeated praise improved pupils' achievement in mechanical tests but not in more complex problem tests;
- Repeated praise improved pupils' achievement on mechanical tests;
- Repeated blame inhibited performance on mechanical tests;
- Intellectually demanding test outcomes are less influenced by either praise or blame;
- Pupils with limited achievements are stimulated to improve their performance by praise but not by blame;
- Pupils with high achievements are stimulated to improve their performance by both praise and blame;
- Anxious pupils' performance improved with praise but was diminished by blame.

2.22 The transatlantic pedagogic strand: real-time observation methods of teaching and learning using coding tables to tally classroom interactions in order to promote the social-emotional climate of classrooms and appraise teachers

2.23 Apart from Stevens (1912) seminal work that employed a systematic method of real classroom observation, subsequent research concerning teacher student interactions used an empirical method and was usually completed by behavioural psychologists.

A distinguishing feature of the pedagogic strand is the absence of behavioural neologisms such as 'reinforcement' used to describe the effect of praise or blame upon pupils.

In the wake of Lewin's developing work on the psychology of groups (1936), there was academic interest from educational researchers in the 'social-emotional climate for learning' in classrooms. Anderson and Brewer (1945; 1946a; 1946b) considered the 'dominative' and 'integrative' aspects of teacher-pupil interactions. They developed a coding schema for quantifying observations of teacher-pupil verbal interactions in the natural teaching environment. The schema categorised 26 teacher-behaviours and 29 child-behaviours that could be separately observed – a complex task for an observer.

2.25 Withall (1949) developed a real-time systematic classroom observation method using a 7 category schema simplified from Anderson and Brewer's work:

> 1. Learner-supportive statements that have the intent of reassuring or commending the pupil.
> 2. Acceptant and clarifying statements having an intent to convey to the pupil the feeling that he was understood and help him elucidate his ideas and feelings.
> 3. Problem-structuring statements or questions which proffer information or raise questions about the problem in an objective manner with intent to facilitate learner's problem solving.
> 4. Neutral statements which comprise polite formalities, administrative comments, verbatim repetition of something that has already been said. No intent inferable.
> 5. Directive or hortative statements with intent to have pupil follow a recommended course of action.
> 6. Reproving or deprecating remarks intended to deter pupil from continued indulgence in present "unacceptable" behavior.
> 7. Teacher self-supporting remarks intended to sustain or justify the teacher's position or course of action. (p. 349)

The scale was used to code transcripts of sound recordings of classes. Inter-observer and test re-test reliability were vouchsafed by Withall. Like Anderson and Brewer, Withall intended the scale to be used to measure and improve the social-emotional climate for learning in the classroom.

2.26 Observational coding systems multiplied over the next 20 years in pedagogic research. Typicallly, Flanders (1964) developed the *Flanders Interaction Analysis System* (FIAC) which was critically favoured by Galton (1978) and by Wragg (1999). To use FIAC, an observer sits in a classroom and every 3 seconds enters a numerical code on a 10 x 10 grid:

Teacher Talk	Indirect Influence	1. Accepts feeling: accepts and clarifies the feelings of the students in a non-threatening manner. Feelings may be positive or negative. Predicting and recalling feelings are included.
		2. Praises or encourages: praises or encourages student action or behaviour. Jokes that release tension, not at the expense of another individual, nodding head or saying 'uh huh?' or 'go on' are included.
		3. Accepts or uses ideas of student: clarifying, building, or developing ideas or suggestions by a student. As teacher brings more of his own ideas into play, shift to category five.
		4. Asks questions: asking a question about content or procedure with the intent that a student may answer.
	Direct Influence	5. Lectures: giving facts or opinions about content or procedures; expressing his own ideas; asking rhetorical questions.
		6. Gives directions: directions, commands, or orders with which a student is expected to comply.
		7. Criticises or justifies authority: statements, intended to change student behaviour from non-acceptable to acceptable pattern, bawling someone out; stating why the teacher is doing what he is doing, extreme self-reference.
Student Talk		8. Student talk - responses: talk by students in response to teacher. Teacher initiates the contact or solicits student statement.
		9. Student talk - initiation: talk by students which they initiate. If 'calling on' student is only to indicate who may talk next, observer must decide whether student wanted to talk. If he did, use this category.
		10. Silence or confusion: pauses, short periods of silence and periods of confusion in which communication cannot be understood by the observer.

Table 1 - Flanders (1964) Interactional Analysis Categories (FIAC)

2.27 Galton (1978) reviewed 41 quantitative and qualitative observational systems in use in the UK at that time. In place of the evaluation of the social-emotional climate of the classroom, Galton proposed a new managerial purpose for systematic classroom observation – teacher appraisal: "The use of observational data either to identify or to validate teaching styles is a small but important development." (p.113)

2.28 Rutter, Maughan, Mortimer, Ouston and Smith's (1979) study of twelve Inner London Education Authority (ILEA) secondary schools, published in book form as the frequently cited and much criticised (e.g. Simpson, 1980; Steller, 1980; Burgess, 1981) *Fifteen Thousand Hours: Secondary Schools and their Effects on Children*, attempted a comprehensive analysis of factors that were indicative of a 'successful' school.

Heath and Clifford (1980) published a fierce critique. They argued that conclusions drawn by Rutter and colleagues were decontextualised and derived using poorly formulated statistical analyses; and that findings had been unjustifiably generalised as applying to all UK secondary schools. They said that this was flawed research that had unmerited influence on both sides of the Atlantic and that Rutter's conclusion: that no underperforming school should be exempted from a requirement for significant school improvement whatever its context and history; and that it was simply a matter of leadership and good management, was simplistic.

Rutter's (1979) study used systematic classroom observation for the purpose of measuring the social climate of the 12 secondary schools and their conduciveness for learning. In total, 402 lessons were observed for 5 minutes at a time, using 5 randomly selected pupils in each to establish a measure of whole-class 'on-task' time. A mean of 81.5% 'on-task' time was extrapolated for the 12 schools. Teachers' verbal behaviour was recorded to establish rates of approval and disapproval. Positive teacher feedback comments were infrequently used with 3 or 4 instances being recorded for each lesson, but higher rates of praise were correlated with better behaviour.

2.29 Wragg (1999) modified FIAC (Flanders, 1964) to develop his own interactional appraisal tool: the *Exeter Schedule* which was used in the Leverhulme Primary Project (Wragg, 1993) which concluded with a list of 'must dos' for teachers in respect of how classroom interaction required their management. Wragg (1999) judged the purpose of systematic classroom observation to be teacher-appraisal and the improvement of teaching: "Skilfully done, classroom observation can be a valuable tool for improving the quality of teaching; badly handled, it can be a menace." (Preface)

2.30 In a constructively critical review, Leary (2014) raised a number of concerns about the objectification of teachers' practice using systematic classroom observation, e.g. by Ofsted, commenting that being able to determine the purpose of such observations is a critical factor in determining the validity and reliability of findings.

2.31 Ofsted regularly publish (e.g. 2015) new guidelines about how schools will be inspected and how classroom observations should be conducted:

> The key objectives of lesson observations are to inform the evaluation of the overall quality of teaching over time and its contribution to learning and achievement, and to

assess the behaviour and safety of pupils and the impact of leadership and management in the classroom. (p.15)

The guidance about how lesson observations will be conducted is not detailed. It is left up to 'professional judgement':

> Lead inspectors should use their professional judgement to plan appropriate observations. For example, inspectors may engage in:
>
> - short visits to a number of lessons,
> - spending a few minutes in each short observations of small group teaching, for example of phonics lesson,
> - observations of more than 25 minutes, during which they may observe activities and talk with pupils about their work tracking a class or specific group of pupils to assess their experience of a school day or part of a school day
> - tracking a class or specific group of pupils to assess their experience of a school day or part of a school day – inspectors may identify a class or classes that contain one or more pupils from the specific groups identified in the pre-inspection analysis. In this way, the experience, progress and learning of these pupils can be judged within the context of other pupils' experience. (p.16)

2.32 The similarity of purpose between 'time and motion' studies in the manufacturing industry (Gilbreth, 1917; Price, 1989) and the classroom observations used by Ofsted for teacher and school appraisal is evident. They share a common ancestry in Thorndike's work at Teachers College, Columbia University (Joncich, 1968).

With time and motion studies in the manufacturing industry and mass production, the optimisation of efficient productivity was the primary goal (Roethlisberger and Dickson, 1939). With the early examples cited of systematic classroom-observation, optimisation or improvement of the social-emotional climate of the classroom was the goal (Anderson and Brewer, 1946; Withall, 1949; Flanders, 1964). A positive social-emotional climate for learning was viewed *a priori* as being a conducive and contributory factor for teaching and learning.

The aim of teacher observation and appraisal according to Ofsted (2015) is ostensibly 'school improvement'. This is measured exclusively by an increase in educational productivity and efficiency, e.g. more A-C grades at GCSE. It is no longer the case that improving the social-emotional climate in a school could be considered a humane 'good-in-

itself' worth striving for by school leadership teams. The work of Anderson and Brewer (1946), Withall (1949) and Flanders (1964) has become redundant.

2.33 The North American (and New Zealand) observational (non-empirical) educational psychology strand: real-time 'naturalistic' time-sampling methods to record classroom interactions

2.34 Olson (1929) of the University of Michigan is credited (Galton, 1978; Skiba, 1989) with inventing the observational method of time-sampling repeating behaviour - for example, the '*nervous habits'* (p.1) of a small sample of children, including: 'tics and grimaces, nose-picking, pulling hair, scratching, sucking thumb rubbing eyes and genital manipulation'. Olson used a schedule of 14 x 10 minute observations over 8 school days with observers filling in a tally sheet.

In a subsequent study, Olson (1931) recorded the frequency of whispering behaviours in an elementary and a high school using a similar method. Olson reported findings thus:

> A group of students in a course in the psychology of personality collected data on the occurrence of whispering in elementary and high school children by the use of a time sampling technique. Each student made ten consecutive five-minute observations. The reliability coefficients for the average of two unpractised observers varied between .33 and .88 in the room studied. No significant sex difference appeared in the mean whisper scores. The amount of whispering present in the various grades appeared to be a function of the situation rather than a reflection of a developmental trend. The evidence suggested that the more intelligent and scholarly whispered somewhat less than other children.' (p. 454)

Olson was concerned about sample-representativeness but was less concerned about inter-observer reliability and specifying confounding contextual variables.

2.35 Florence Goodenough (1928, 1930) of the University of Minnesota is also credited (O'Connell, 1990; Pickren, Dewsbury and Wertheimer, 2012) with inventing the observation method of time-sampling. Goodenough used 5 sets of observers to measure the frequency of 5 specified behaviours for short periods on succeeding days. Goodenough (1930) commented:

> In a small number of instances it was necessary to make two observations of a child on the same day in order to make up for losses due to absentees. When this was done,

> the two observations were always separated by as great a time interval as possible. As a rule, one was taken in the morning, the other in the afternoon. In this respect, the procedure differed from that used by W.C. Olson. (p. 29)

Subsequently, Goodenough (1937) criticised her own and Olson's observational time-sampling method in favour of an experimental methodology. Goodenough concluded that it had been impossible to isolate the effect of confounding variables such as observer-effects and environmental factors within classroom settings.

2.36 Neither Goodenough (1928) nor Olson (1929) were informed sufficiently to differentiate between alternative time-sampling methods and tendencies towards bias that different methods risked.

Previously, there had been two types of time-based behaviour recording schedules used in time and motion studies in the manufacturing industry (Price, 1989). These schedules did not employ periodic-interval time-sampling:

1. Frequency-counting - how many times does a specified behaviour or event occur within an observed period, e.g. a lesson;
2. Duration - how long does a specified behaviour persist for.

These two methods appeared to work well-enough for reporting simple idiographic data but are less amenable to more complex comparisons of behaviours, correlations of behaviours, or when making generalisations about behaviours in order to support or challenge a psychological hypothesis.

2.37 Three main types of interval-recording time-sampling protocols are identified by Skiba (1989):

1. Momentary time-sampling (MTS), e.g. every 2 seconds observing subject momentarily to see whether they are scratching head (or not);
2. Whole interval time-sampling (WITS), e.g. 2 second observation every 60 seconds – if scratching head for whole of 2 seconds, tick 'yes';
3. Partial interval time-sampling (PITS), e.g. at least one instance of a facial tic within a 2 seconds time-slice, every 60 seconds.

Additionally, Bart, Fligner and Notz. (1998) described Periodic-Interval Event Sampling (PIES), e.g. each day at a set time for 2 minutes, tally how many times the dog barks.

2.38 Extrapolating a percentage-occurrence of a specified behaviour in a long time-period, it can be predicted that PITS and WITS might over-estimate the frequency. Millar and Hawkins (1976) reported that PITS often over-estimated frequency by between 49% and 89%. Powell, Martindale, Kulp, Martindale, and Bauman. (1977) concluded in favour of the accuracy of MTS. Steffen (1985) concluded similarly:

> In summary: the momentary time-sampling recording method always gave closer estimations to the session [total] response duration than the whole and the partial interval recording methods. The accuracy of all three methods to measure the session response duration were enhanced when short and frequent recording intervals were used. (p. 32)

Steffen's conclusion was supported by Skiba's comprehensive review (1989).

2.39 Professor Mary White of the Teachers College, Columbia, and Director of the 'Center for Behavioral Analysis of School Learning' is credited by Swinson and Harrop[11] (2012) with completing the first study (White, 1975) of the association between teacher approval and disapproval, and the behaviour of a class of students, using a time-sampling method. White organised 16 classroom observations in 16 schools; Grade 1 to Grade 12. Inter-observer reliability was tested and reported. White's findings supported the hypothesis that teachers used less approval and more disapproval as grade increased:

> 1. Pupils received more total teacher disapproval than total teacher approval over grade;
> 2. for instructional behaviour alone, teacher approval rate was higher than teacher disapproval rate, a discrepancy that was particularly marked in the primary grades;
> 3. for managerial behavior alone, teacher disapproval far outweighed teacher approval. Teacher approval for managerial behavior was almost nonexistent. Pupils almost never heard a teacher say such things as: "How nice, you are on time!", "I like the way you are sitting", "You are behaving so well!" (p.370)

White used the language of behavioural psychology: "How do rates of teacher verbal reinforcement operate to maintain or increase the school behavior of pupils?" (p.367)

[11] A time-sampling study by Madsen *et al.* (1968) utilised a simple 10 second time-slice unit to observe and categorise the behaviour of 2 kindergarten students in 2 classrooms after a behavioural intervention had been used with them but the research was not intended as a systematic whole-classroom observation.

White's term: 'instructional', referred to student behaviour that was identified as academic in purpose; 'managerial', referred to student behaviour that was social, anti-social or organisational in purpose.

White (1975) suggested the following directions for future research:

> 1. Teachers find immediate disapproval of a managerial behavior more reinforcing for them as teachers than ignoring the behavior, or approving an alternate behavior.
> 2. A Law of Personal Effectiveness may be operating that would explain hypothesis one above, namely: teachers (and others) are highly rewarded by those behaviors of their own that have an immediate effect upon the environment.
> 3. Rates of approval and disapproval received by teachers from the school environment influence their emitted rates of approval and disapproval.
> 4. The slower the rate of demonstrated pupil learning, the lower will be the rate of teacher approval; the faster the rate of demonstrated pupil learning, the higher the rate of teacher approval.
> 5. Increasing the rates of teacher approval for instructional and for managerial behaviors in junior and senior high school will increase the rate of learning and of appropriate behaviors in such students.
> 6. At the high-school level, student learning rates for social behaviors exceed the rates for academic behaviors. This is due to the higher rates of approval emitted by peers for social behaviors, as compared to the lower rates of teacher approval for instructional behaviors. (p.371)

2.40 White's was a seminal study. It was the first that time-sampled natural frequencies of teacher approval and disapproval in the classroom. It also provided inspiration for a stream of British educational psychology research (Merrett and Wheldall, 1986; Wheldall, Houghton and Merrett., 1989; McNamara and Jolly, 1990; Harrop and Swinson, 2000; Swinson, 2005; etc.), including the author's previous, (Apter, *et al.*, 2010) and current research. As noted by Swinson and Harrop (2012), White's findings went largely unchallenged during the 1980s, in North America and in the UK.

2.41 White (1975) designed the *Teacher Approval and Disapproval Observation Record* (TAD). This schedule focussed the observers' attention during a 20 second timed period on a teacher's verbal behaviour, transcribing their approving comments and their disapproving comments. The observer also described the student behaviour that had elicited the response.

White detailed what was meant by approval and disapproval:

> Teacher approval was defined as a verbal praise or encouragement, and teacher disapproval as a verbal criticism, reproach or a statement that indicated that the student's behavior should change from what was unacceptable to acceptable by the teacher… …Teacher verbal responses that did not fall into either of the above categories were regarded as neutral and were not recorded. (p.368)

Evident methodological problems were:
1. Post-hoc description of behaviour associated with the verbal response of the teacher is open to observer-interpretation error.
2. 20 second time allowance for recording means that fast subsequent exchanges are lost whilst the observer records first interaction.
3. Using a time-slice unit of 20 seconds with a multiplier to produce an estimation of total approval / disapproval for the observation period must have produced over-estimations of both (as predicted by Milar and Hawkins, 1976).
4. There was no device in the design for determining how much (duration) 'on-task' and 'off-task' behaviour occurred during a particular observation.

2.42 A year later on the West Coast of North America, Persons, Brassel and Rollins. (1976) developed: "A practical observation procedure for monitoring four behaviors relevant to classroom management." The four behaviors were:
A. a teacher's use of 'positive events', including: the use of verbal praise, the granting of privileges, positive physical contact and consumables / tokens (rewards);
B. a teacher's use of 'negative events', including: the use of verbal criticism or sarcasm, withdrawing privileges, isolation, aversive physical contact - such as: 'spanking', 'slapping', 'shaking' and 'grabbing';
C. student disruption, including: talking, out of seat without permission, generating loud noises, disturbing other students verbally or physically, or by interfering with their possessions;
D. student attention: student-time, attending to a teacher-directed task in seconds.

The repeating observational sequence of 15 minutes entailed the observer first focussing on the teacher for 5 minutes in order to record tallies of instances of A. and B. behaviours. They then switched their attention to a third of the class at a time, spending 5 minutes recording C. as a tally of separate disruptive incidents, followed by 5 minutes recording D. attentive behaviour. The attentiveness of each student was observed for 20 seconds until

each student in the observation-group had been observed. Attentiveness was graded as 'involved' (0-5 seconds off-task), 'medium involved' (6-15 seconds off-task) or 'uninvolved' (16-20 seconds off-task).

During the next repetition the observer concentrated on the next third of the class, and on the final repetition, on the last third. Three repetitions of the sequence took 45 minutes in total, during which the teacher would have been observed for the first 15 minutes of the lesson. This was unlikely to be representative of the teacher's verbal behaviour during the last 15 minutes of the lesson.

2.43 It was intended that a correlated association could be made between teacher verbal behaviour and student learning behaviour. The results using Pearson's product moment coefficients were tabulated thus:

	Teacher behavior		*Student behavior*	
Behavior Observed	*Positive events*	*Negative events*	*Disruption*	*Attention*
Teacher positive events	-	*.33*	*.39*	*.66***
Teacher negative events	*.28*	-	*.48***	*.22*
Student disruption	*.45**	*.60***	-	*.61***
Student attention	*.58***	*.33*	*.76***	-

Note: *p < .05; **p < .01. Coefficients in the upper right hand section are based on observations made in September; coefficients in the lower left hand section are based on observations made in April and May'

Table 2 – Inter-correlations of student and teacher behavior (Persons, et al. 1976, p.70)

Persons *et al.* comment:

> At both the beginning and the end of the school year, the teacher's use of positive events is related to student disruption in a negative direction and to student attention in a positive direction. On the other hand, the teacher's use of negative events is correlated positively with disruptive student behavior. The direction of these correlations conforms with common sense expectations. (p.70)

2.44 The method that Persons and colleagues developed and the results that they obtained was arguably more influential with UK researchers (Wheldall, Merrett and Borg, 1985; Merrett and Wheldall, 1987; Wheldall, *et al.*, 1989) than White (1975).

Persons. *et al.*'s observation schedule (1976) was similar to Merrett, Wheldall and Borg's (1985) observation schedule: 'Class Teacher / Pupil Observation Schedule CTP2'.

The method could be described as *mixed-interval*: a combination of *partial-interval* observations of a sample of students' behaviours, followed by a *partial-interval* tally of specified teacher verbal behaviours.

Unlike Wheldall and his Birmingham (UK) colleagues, Persons *et al.* (1976) appeared to distance their study from the behavioural paradigm and did not use behaviourist terminology such as 'reinforcement'.

2.45 In New Zealand, Ted Glynn and colleagues (Thomas, Presland, Grant and Glynn, 1978) conducted a small scale study using a time-sample schedule to record the approval and disapproval rates of 10 teachers in 3 intermediate (junior-high / middle) schools and associations with 'on-task' and 'off-task' behaviours. The schedule used was a modification of White's (1975) method, and randomly identified 10 children in each class to be observed, who would be statistically representative of the class. They adopted a partial-interval time sampling protocol: if a particular student was 'off-task' at all during a 10 second timed period, they were marked as being 'off-task' for that interval.

During the same period, the observer was directed to mark down instances of teacher verbal approval and disapproval. Variable figures for 'on-task' behaviour were obtained: between 43% and 90%. It was presumably difficult to manage an accurate observation of the teacher's verbal behaviour. The method would also be expected to over-estimate 'off-task' (Steffen, 1985).

2.46 Nafpaktitus, Meyer and Butterworth (1985; California State University) developed an observation method that cited White's findings (1975) and attempted to address some of the methodological problems. They noted that researchers such as Thomas, *et al.* (1978) who had adhered to White's method had not succeeded in reporting 'naturally occurring' (p.362) rates of teacher approval and disapproval. The study was ambitious with 29 schools taking part and teachers being observed in each school 3 times to determine rates of approval and disapproval associated with 'on-task' behaviour.

A randomly-selected sample of 6 students in each class were observation for 10 seconds-each to establish whether each student was 'attending'; 'on-task'; or 'off-task'. The same 6

students were observed in 3 different academic lessons with the same teacher but only the last observation was included in the results (to reduce any observer effects).

After each of the 6 students had been observed for 10 seconds and their behaviour recorded on the record sheet - a further 5 seconds was allowed for this for each student – the observer then changed the focus of their observation to the teacher.

For the next 5 minutes, the teacher's verbal behaviour was observed, categorised and recorded. Teachers' verbal behaviour was categorised as 'appropriate approval', 'disapproval' or 'inappropriate approval'. Any verbal behaviour that did not fit into these categories was ignored and not recorded.

2.47 As with Persons *et al.* (1976) it was irrelevant to the purpose of Nafpaktitus, *et al.*'s observational method which student the teacher was interacting with: one of the chosen 6 students or any other student in the class. Two important experimental assumptions were being made that: the 6 observed students' behaviour were indicative of the average 'class behaviour' during that lesson with that teacher; and the snapshot of the teacher's behaviour was indicative of that teacher's usual verbal behaviour with that class, during that lesson.

Each one of the 3 class-observations involving a particular teacher lasted 30 minutes in total, of which 15 minutes was teacher-focussed and 15 student-focussed in alternating 5 minute segments.

The researchers found (ibid.):
1. Higher rates of inappropriate approval of 'off-task' behaviour were *associated* with lower rates of 'on-task' behaviour.
2. The lowest rates of teacher disapproval were associated with the highest rates of student 'on-task' behaviour.
3. Mean rates of teacher approval were found to exceed rates of disapproval.

Like Persons *et al*. (1976) report, Nafpaktitus *et al.* (1985) distanced their findings from the behavioural paradigm using 'associated with…' rather than the causal behavioural verb: 'reinforced'. The emergence of a dynamic interactional model was evident: "In the feedback system of the classroom, students continually influence teacher behavior and vice versa." (p. 366)

Extrapolating from the tabulated results (Table 3; p.365) the mean rate of 'on-task' behaviour recorded during this study was 71.7% (+ or - 2.5%). The authors commented: "In conclusion, our interpretation of the data suggests that to achieve an acceptable level of on-task behavior, say approximately 80% or better, teachers must provide high rates of appropriate approval combined with both low rates of inappropriate approval and disapproval." (p.367)

2.48 The UK observational (non-empirical) educational psychology strand: real-time 'naturalistic' time-sampling methods to record classroom interactions

2.49 Kevin Wheldall and colleagues of the University of Birmingham's educational psychology professional training course published research (Wheldall, Merrett and Borg, 1985) using the *Class Teacher / Pupil Observation Schedule* (CTP2) to evaluate the *Behavioural Approach To Teaching Package* (BATPACK) that they had developed. The CTP2 was used to evaluate classroom behaviour before and after the BATPACK intervention had been employed in two primary schools.

The CTP2 was developed further by Frank Merrett and it was renamed the 'Observing Pupils and Teachers in Classrooms' schedule – OPTIC, (Merrett and Wheldall, 1986). Merrett and Wheldall proposed that: "Because of the extreme complexity of the interaction taking place in a classroom, it is impossible to achieve an objective description of the totality of what is happening there." (p.57)

Cited by Merrett and Wheldall, the point is expanded in a corollary by McIntyre (1980):

> Only by recognising that he [sic] must ignore much that is happening and by focussing on carefully selected and predefined facets of classroom activity can the observer hope to avoid a subjectivity, of which he cannot himself be aware, and to provide descriptive evidence in which he and others can place some confidence. It is, then, a tradition which aspires to objectivity of observation. It seeks to achieve this through abstracting from the complexity of classroom life aspects of it which are hypothesised to be of particular significance in relation to some given question, and through the selective observation and categorisation of these relevant aspects according to a predetermined set of rules. (p.3)

2.50 As with North American protocols (White, 1975; Persons *et al.*, 1976), OPTIC (Merrett and Wheldall, 1986) required an observer to complete observational cycles using alternating procedures, focussing in turn on teacher behaviour and pupil behaviour, thus:

A. Observation of teacher: There was to be three minute observations on five occasions, interspersed with the observations of students (see B. below). Events were categorised as verbal praise, or criticism. The observer recorded whether the verbal comment is directed towards the academic or social behaviour of a student. Criticism also included: "Gestural responses like frowning or glaring; aversive contact involving shaking or smacking; withdrawal of privileges; isolation from the rest of the group, and so on." (p.68)

   The OPTIC schedule also required the observer to annotate their record as to whether they judged the 'event': appropriately contingent or not; related to a rule; used as an exemplar to the rest of the class; or the contingency explained and noted to the class, e.g. "Well done! That was very neat work." (p.69)

B. Observation of students: "The observer should divide the class by eye into three convenient and approximately equal groups and pay attention to each in turn for one minute. During the first minute the observer should look at each pupil in the first group, in turn, for four seconds and decide for the whole of that period he / she is on-task. To be rated on-task the child must be attending, i.e. in eye-contact with the teacher or the task or otherwise following the teacher's instructions. This would include listening to the teacher or to some apparatus operated by the teacher or to another child asked by the teacher to speak." (p.69)

Each classroom observation usually took 40 minutes. When timings were adhered to, the teacher would have been observed for fifteen minutes in total.

2.51 There were a number of methodological problems with OPTIC. Accepting McIntyre's (1980) corollary, it is nonetheless debatable as to whether the variables that OPTIC restricted itself to measuring were precisely and robustly defined. There were also vagueness about the descriptors of the teacher's verbal behaviour and reliance upon the observer's judgement as to what was 'appropriately contingent'.

Potential criticisms were partly addressed by testing OPTIC's inter-observer reliability. Using a percentage-agreement rubric, 'two experienced observers' each observed five

teachers using an OPTIC schedule and achieved 93% agreement for teacher behaviour and 94% for student 'on-task' behaviour[12].

The observation of students followed a PTI protocol. If a student's attention was interrupted momentarily during a 4 second observation, they would be recorded as being 'off-task' for that whole 4 second period. This means that a student's 'on-task' behaviour would probably be under-estimated (Steffen, 1985).

2.52 The similarity overall of OPTIC with the method of Persons *et al.* (1976), developed some 10 years previously, is also worth considering in respect of the differences and whether Merrett and Wheldall's developments were improvements.

Persons *et al.* began each round of observations by observing the teacher's interactions – but for five minutes in contrast to OPTIC's three. This was followed by a five minute observation of one of three approximately equal groups to tally 'disruptive' incidents; and then an observation of each student in each group for twenty seconds each. Persons *et al.*'s observers made their own judgement as to what the student's most *frequent* behaviour was ('on-task' or 'off-task') during a twenty second observation; compared with Merrett and Wheldall's PTI (partial-time interval) judgement within a four second observation.

Merrett and Wheldall dispensed with a separate observation of 'disruptive behaviour'. This seems sensible as Persons *et al.*'s method depended considerably on the observer's judgement as to what was 'disruptive behaviour' and what was not.

2.53 Merrett and Wheldall (1986) cited White (1975) and Persons *et al.* (1976) in describing the development of OPTIC; but did not use Glynn's (Thomas, *et al.*, 1978) innovation, subsequently adopted by Nafpaktitus *et al.*, (1985), of only observing a proportion of randomly selected students in a class as a representative sample. Wheldall and Merrett's methodological decision to observe every student for four seconds each meant that teachers were only observed for fifteen minutes during a 40 minute observation. OPTIC was thus less focused upon, and less accurate about teacher's verbal behaviour.

[12] It is not stated that the 'two experienced observers' were not Merrett and Wheldall themselves however.

2.54 A year later, Merrett and Wheldall (1987) published a study of 128 primary and middle school teachers using the OPTIC schedule. The language used in Merrett and Wheldall's research report is psychological and behavioural rather than pedagogical. They describe teachers using: "social reinforcement" (p.95), "very lean intermittent reinforcement schedules" (p.95) and "tending to emit lower rates of approval and disapproval than their American counterparts" (p.96) when describing teachers verbally approving or disapproving of pupils' work or behaviour.

2.55 Each teacher was observed for either four 9 or 15 minutes sessions. Because these were primary and middle schools, the same teacher and class were observed working together. Depending on when the observation occurred in the 4 observed lessons, e.g. during the first or last 15 minutes of the lesson, different teacher and pupil behaviours would be expected. The first part of an academic lesson in a primary school is usually characterised by a teacher's verbal exposition and task-setting. Pupils behave in a way that is particular to this type of teacher behaviour during this time. If all 15 minute observations of the 128 teachers were conducted at the beginning of lessons, we might expect teachers' talk and pupils' listening to be measured at higher rates than at other times during lessons. This would be predicted to significantly affect observations of pupils' 'on-task' and 'off-task' behaviour.

2.56 Merrett and Wheldall's headline finding was teachers used more approval than disapproval, but mainly directed at *academic* behaviour. In contrast, children were more often criticised for *social* behaviour than praised. Results were expressed in the following table:

Behaviour	*Approval*	*Disapproval*	Total
Academic	50%	16%	66%
Social	6%	28%	34%
Total	56%	44%	100%

Table 3 – Primary and middle school teachers' use of approval and disapproval (Merrett and Wheldall, 1987, p.97)

That teachers used 3 times more approval for academic behaviour and 3 times more disapproval for social behaviour became an established research finding, replicated by

Wheldall and Birmingham colleagues, others in the UK, (Wheldall, *et al.*, 1989; McNamara and Jolly, 1992; Merrett and Wheldall, 1992; Swinson and Harrup, 2001; Apter, *et al.*,2010) and internationally (Akin-Little, Eckert, Lovett and Little., 2004; Sazak Pinar and Güner Yildiz, 2013; Güner Yildiz and Sazak Pinar, 2014).

Merrett and Wheldall (1987) ruled out the effect of a number of contextual variables: the sex, age of teachers and age of pupils, by claiming that effects had balanced out in the sample, e.g., equal numbers of male and female teachers and pupils in the sample, but they write that teachers of the youngest age cohorts were entirely female. From the 5 year-old grouping upwards, there were progressively more positive comments used by teachers up to the 12 year-old age grouping. Merrett and Wheldall reported that this finding was dependent upon pupil age and not teacher gender.

Whilst there was significant variability in teacher responses to student 'on-task' and 'off-task' behaviour, Merrett and Wheldall reported that their findings over-turned Persons *et al.*'s (1976) finding that teachers' positive comments were associated with 'on-task' behaviour: "We found very little evidence for this. The strongest predictor[13] of on-task behaviour was a negative correlation with disapproval for social behaviour." (p.101)

2.57 Following their primary study, Wheldall, Houghton and Merrett, (1989) used OPTIC to observe 130 secondary school teachers in one lesson each. Class size varied between 8 and 32, included all secondary age-groups, and non-academic classes were excluded from the study. The findings are shown in the following table:

Behaviour	*Approval*	*Disapproval*	Total
Academic	45%	15%	60%
Social	10%	30%	40%
Total	55%	45%	100%

Table 4 – Secondary teachers' use of approval and disapproval (Wheldall, Houghton and Merrett, 1989, p.42)

[13] The wording suggests that if teachers verbally managed behaviour more, pupils would stay on-task, but a correlation does not 'predict' this (Baguley, 2010). The causal vector might operate in the opposite direction. 'Off-task' behaviour of students might cause teachers to use more critical verbal comments directed at student behaviour.

Again, 3:1 academic-approval to disapproval and 1:3 social-behaviour approval to disapproval ratios were revealed. Lessons of teachers who used more approval in respect of academic behaviour and social behaviour and less disapproval of social behaviour were associated with higher levels of 'on-task' behaviour. An additional finding was that female teachers disapproved of social behaviour significantly more frequently than male colleagues ($p < 0.02$).

The mean rate of 'on-task' behaviour was 80.5% (SD= 15.5). Rates varied between 17% and 99%:

> The highest mean level of on-task behaviour was recorded in the first year [Y7] classes where pupils were getting on with their work for approximately 88.7% of the time. From year one, however, levels of on-task behaviour fell sharply reaching 74.6% in year three [Y9] before rising again to over 80% in years four and five [Y11 and Y12]. (p.44)

Wheldall *et al.* (1989) reported associated differences between rates of approval, disapproval and 'on-task' behaviour, by subject, recording the highest rate of 88.6% for maths and the lowest (excepting a miscellaneous category) of 76.3% for modern languages.

In a follow-up secondary phase study (Houghton, Wheldall, Jukes and Sharpe, 1990) using OPTIC, Wheldall and colleagues demonstrated an increase in 'on-task' classroom behaviour from a baseline of 69% of 13 % to 82%. This increase was associated by the authors with teachers using private rather than public reprimands.

2.58 In another OPTIC study (Merrett and Wheldall, 1992) observed classes in both primary and secondary schools. The researchers noted no differences in the primary sample but significant differences in the secondary sample, between the way in which teachers praised or reprimanded pupils of each gender: "For the sample of secondary teachers, however, there was evidence for major significant differences in rates of responding to boys and girls, boys receiving more responses overall (both positive and negative) from teachers" (p. 73). The mean 'on-task' behaviour recorded for the primary sample was 73.5% and 74.5% for the secondary sample.

2.59 Merrett and Wheldall's schools used in the research cited were based in the UK locale of Walsall, in the West Midlands. The 'tawse' – a leather spanking implement was still used

for punishment by teachers in Walsall schools (Walsall Observer, 1975). Most schools in London and the south had dispensed with corporal punishment (Farrell, 2015). Corporal punishment in all UK schools was outlawed by the UK Government in 1987 (HM Government, 1987). Whilst Wheldall and Merrett were most interested in verbal criticism, findings were probably confounded by the implicit threat of corporal punishment in the schools they used. What this meant in respect of 'on-task' rates reported by Wheldall and colleagues in the 1980s is a matter for speculation.

2.60 A distinctive classroom observation schedule was developed by Jolly and McNamara (1992). McNamara's schedule focussed more observer-time on pupils by requiring observers to momentarily time-sample (MTS) every pupil in a class as 'on-task' or not. The schedule was suited to exploring changes in class behaviour after an intervention or contextual change.

McNamara and Jolly (ibid.) observed the behaviour of a class of 15 pupils – described as a 'disruptive class' (p.240) - in an R.E. lesson and in a French Lesson. They were observed again after the teacher has been coached in verbally promoting 'on-task' behaviour. 'on-task' behaviour was shown to have significantly increased and disruptive behaviour to have decreased.

A criticism of this 'ABA' (repeated baseline-measure) experimental design is that other contextual variables, e.g. time of day, subject matter, fatigue, observer effects, are difficult to isolate. A strength is the accurate detail it can reveal in respect of pupil variables, e.g. gender and seating position.

2.61 Reporting an investigation of the use of approval and disapproval by 10 infant, 10 primary and 10 secondary school teachers, Harrop and Swinson (2000) used McNamara's MTS schedule and tape-recorded the verbal interactions of teachers with pupils. Independent scrutinisers categorised teachers' comments from transcripts and established ratios between approval and disapproval in different settings.

Harrop and Swinson's (2000) findings replicated findings from other studies that approval was given primarily to academic rather than social behaviours, with the reverse being the case for disapproval. However, an important caveat was reported: "It is worth noting, however that the data [in the table] show that the secondary teachers, unlike the infant and

junior teachers, actually gave more disapproval to academic behaviours than to social behaviours." (p.479)

In subsequent research, Swinson and Harrop (2001) used the same method in ten junior and ten infant classrooms. 'on-task' behaviour was again associated with teacher verbal approval but a difference was found in respect of the curvilinear relationship that the researchers found between disapproval and 'on-task' behaviour: "It is tempting to conclude that disapproval is effective in maintaining on-task behaviour provided that it is not used too much, so that beyond a certain level it becomes counter-productive." (p. 164)

It is understood from the authors' that an accurate correspondence of the MTS data-grid with the transcription of teachers' verbal behaviour and a predictive relationship between teacher verbal feedback and particular instances of student behaviour was not established.

2.62 Adopting Persons *et al.* (1976) method, Apter [author] (Apter, *et al.* 2010) added Nafpaktitus, *et al.*'s (1985) adaptation of only observing a small random sample of students in a class in order to obtain an estimation of a class's 'on-task' behaviour. Apter's *Teaching and Managing Behaviour in School: Time Sample Record* (*TaMBiS TSR*[14]) required an observer to randomly select three students in a primary classroom, and momentarily observe them in-turn to record whether they were 'On-task'[15] or not.

After each student-observation (MTS) sequence, the observer recorded the teacher's verbal behaviour for a specified time, e.g. 2 minutes. The teacher's verbal behaviour was tallied thus:
'ARD': Academic Routine Directions;
'TAP': Academic Positive comments;
'TAN': Academic Negative comments;
'TSP': Social Positive comments;
'TSN': Social Negative comments.

The 'student MTS / teacher tally' cycle would be repeated throughout an observation.

[14] An alternative function of the TaMBiS TSR was to compare the 'On-task' behaviour of a student, who had been identified as challenging by the teacher with two randomly selected student comparators in a class. The schedule was originally developed for trainee-psychologists to use in Wolverhampton (UK) schools to support them in advising teachers about managing non-compliant behaviour.
[15] Variable names that were designated by the author are displayed in the text within single quotation marks, e.g. 'On-task' and 'ARD'.

McBer (2000) had proposed that verbal pace was an important feature of effective teaching:

> The teachers presented information to the pupils with a high degree of clarity and enthusiasm and, when giving basic instruction, the lessons proceeded at a brisk pace… However, it was evident that when the effective teachers were not actively leading the instructions they were always on the move, monitoring pupils' focus and understanding of materials. (p.36)

Apter introduced the 'ARD' category in *TaMBiS TSR* following piloting of the materials in Wolverhampton in 2004. The 'ARD' category was intended to capture teacher-talk that was neither approving nor disapproving was a dominant feature of classroom discourse in primary schools that could be usefully captured.

The *TaMBiS TSR* was tested for inter-observer reliability, calculating the statistical coefficient *kappa* (Cohen, 1960) at between 0.77 and 0.89 for joint observations where a value in excess of 0.75 indicated significant inter-observer reliability.

Across the UK, 141 lessons in primary schools were observed by 71 psychologist-observers. Each psychologist-observer usually returned 2 sets of observations - one from a morning lesson and one from an afternoon lesson. Apter, *et al.*'s (2010) was the largest observational study of naturally occurring classroom interactions at that time.

Mean 'On-task' student behaviour was reported as 85.11% - significantly higher than previous studies had revealed. The highest correlations were between students' 'On-task' behaviour and teachers' positive academic feedback ('TAP'; $r = .345$, $p < .01$); and between 'On-task' and teachers' academic routine directions ('ARD'; $r = .168$, $p < .05$).

Ratios of teacher-feedback comments were different to those previously found: 5 to 1, academic-approval to disapproval; and 1 to 2, social-behaviour approval to disapproval:

	Positive	*Negative*
Academic	46.2%	9.1%
Social	14.8%	30%

Table 5 – The percentage of positive and negative verbal feedback directed by teachers towards students' academic work and towards their social behavior (Apter, *et al.* 2010, p. 156)

Students were not more 'on-task' to a significantly higher degree ($p<0.05$) in the prescribed lessons of the Literacy and Numeracy Strategies (DfE, 2011) despite teachers using more verbal directions ('ARD') in comparison to non-strategy lessons; but teachers used more positive comments about social behaviour ('TSP') than previously reported, e.g. by Harrop and Swinson, (2000); and Merrett and Wheldall, (1987).

2.63 The table shows the mean 'on-task' classroom-behaviour since 1978 by school-phase and year from systematic classroom-observation studies cited for this review:

Study	Year	Years / phase / country	No. of lessons observed	Mean 'on-task' percent.	Standard deviation
Thomas *et al*.	1978	Y7 NZ	36	66	NR
Rutter *et al.*	1979	Y7-Y12 Secondary UK	402	81.5	NR
Nafpaktitus *et al.*	1985	Y6 – Y9 Senior USA	84	71.7*	NR
Merrett & Wheldall	1987	Y1-Y6 Primary UK	128	69	12.8
Wheldall *et al.*	1989	Y7-Y12 Secondary UK	130	80.5	15.5
Houghton et al.	1990	Y9 Secondary UK	4	69	NR
Merrett & Wheldall	1992	Y3–Y6 Primary UK	32	73.5*	10.5*
Merrett & Wheldall	1992	Y7–Y11 Secondary UK	38	74.5*	13.6*
Swinson and Harrop	2001	Y1–Y2 Infant	16	81.2	7.5
Swinson and Harrop	2001	Y3–Y6 Junior	16	78.5	13.6
Swinson and Harrop	2005	Y7–Y11 Secondary UK	18	80.5	13.1
Apter *et al.*	2010	Y1-Y6 Primary UK	141	85.11	12.1

Key: NR=Not Reported; * Figure extrapolated from data tabulated in alternative form in original text

Table 6 - Quantitative studies where a dependent variable was 'on-task' behaviour expressed as a mean average

2.64 The significance of the research cited here is attributed because of associations made between types of teachers' verbal feedback and student 'on-task' behaviour, revealed as a result of systematic quantitative observational whole-class research.

McNamara and Jolly (1990) argued that whilst 'off-task' does not explicitly include intensely disruptive behaviour, increasing 'on-task' has a pervasive effect: "When disruptive classroom is dealt with by the promotion of 'on-task' behaviours, the total amount of all types of 'off-task' behaviours, from innocuous to grossly disruptive, is reduced." (p.239)

'On-task' behaviour and 'off-task' behaviour are binarily observationally exclusive. 'Off-task' behaviour therefore includes disruptive behaviour. Researchers since Persons *et al.* (1976) have avoided distinguishing between 'disruptive behaviour' and 'off-task behaviour'.

2.65 Whether causality can be claimed between a teacher's positive feedback and an increasing percentage of 'on-task' academic behaviour used by students is a question judiciously side-stepped by researchers. As Nafpaktitus, *et al.* (1985) comment: "Although it is not possible to draw cause and effect conclusions from a correlational non-experimental study, the relations [associations] reported here can be interpreted in consonance with the literature on manipulated rates of teacher approval." (p.365)

Wheldall, *et al*. (1989) choose their words carefully:

> These strong correlations suggest that teachers who use more approval to academic behaviour and social behaviour and less disapproval to social behaviour experience higher levels of 'on-task' behaviour in their classrooms. It must be recognised, however, that no firm conclusions about causality can be reached since these results are purely correlational. (p.45)

The behaviourally causal relationship between stimulus-event and response-event is of considerable interest but is difficult to establish.

To provide an example, if in a particular classroom, a teacher uses a high level of negative feedback directed towards behaviour, and simultaneously, there is a high level of anti-social behaviour, a statistician will either: determine an 'association' exists without establishing causality; or they will assume the direction of causality by determining one type of event:

perhaps the teacher's verbal behaviour, to be the 'dependent variable', and the anti-social behaviour to be the 'independent variable'. They might then use statistical regression to confirm the assumption.

Conversely, the statistician might assume the opposite: that the anti-social behaviour is the 'dependent variable' that is apparently *caused* by the teacher's verbal behaviour: the 'independent variable'. Statistical regression would appear to confirm this assumption as well.

2.66 Accepting McIntyre's (1980) caveat that the complexity of interactive variables operating within the classroom is impossible to comprehensively and accurately capture by any observational research method, it seems useful to try to pragmatically resolve this vagary. One way would be to define and specify a more general concept of a 'social-emotional climate for learning' as developed previously by Anderson and Brewer (1946a, 1946b) and Withal (1949). The 'conduciveness' of this climate would be indicated by a set of quantitative measures, e.g. the amount and type of teacher-feedback that is used; and the amount of 'on-task' behaviour that is observed. There might be other indicators.

Developing the social-emotional climate for learning is a complex aspiration, the wisdom of which is probably dependent upon a researcher's pedagogical standpoint and the weight they give to different psychological paradigms and associated theories of teaching and learning, e.g. behaviourism (Watson, 1913; Skinner, 1938); or social learning theory (Bandura, 1977); or constructivism (Dewey, 1902; Piaget, 1954; Bruner 1996).

Whichever psychological camp a researcher inhabits, the aspiration of using systematic classroom observation to measure, index and improve the social-emotional climate for learning in the classroom is motivated by a different purpose than teacher appraisal.

Using systematic classroom observation to promote the social-emotional climate for learning could be argued to contribute generally to the good of society; not only by promoting humane and productive teaching and learning, but also by promoting students' and their teachers' personal and social development, creativity, compassion, empathy and citizenship, by monitoring, critiquing and improving a set of simple quantitative indicators.

2.67 Research questions and hypotheses

The following research questions and hypotheses were germinated by the initial curiosity described in the introduction (Chapter 1). The relationship between the literature review; and the research questions and hypotheses, does not have a one to one thematic correspondence however. There are expedient omissions in the literature review. For example, the literature review does not consider previous research relating to the contextual variables: 'Locality type'; 'Free school meals %'; 'School roll'; 'Year group'; 'Teacher gender'; 'Years [of teaching] experience'; 'Subject'; 'Adults present [in class]'; 'Class size'; and whether lessons were taught 'AM or PM'. An investigation of these variables was included in the author's previous primary school research (Apter, *et al.* 2010) and the investigation of a similar set of contextual variables was included here in order to provide an analogue for interested readers between the current secondary school findings and those previous primary school findings.

In contrast, an investigation of teacher verbal feedback and significant associations with students' following teachers' directions (or not) was a primary theme of both the literature review and of the current secondary school study. Another primary theme of both review and study was the development of the sampling method used for the systematic observation of the classroom. These two primary themes are significantly represented in the research questions and hypotheses that drive the current research and which are listed here.

2.68 Research question #1: What types of teachers' verbal feedback encourage students to follow their directions and stay 'on-task' in academic lessons in UK secondary schools?

2.69 Hypothesis #1: The time spent by students following a teacher's directions as measured using a quantified dependent variable ('On-task%) is causally associated with higher frequencies of positive verbal feedback by teachers in academic lessons in UK secondary schools directed towards their productive academic-task-focussed behaviour and product, e.g. written work.

2.70 Hypothesis #2: High levels of time following a teacher's directions in secondary classrooms ('On-task%') are associated with teachers who use more verbal teaching behaviour, i.e. students are more 'on-task' when teachers talk more.

2.71 Hypothesis #3: Teachers in secondary schools use significantly more positive verbal feedback, e.g. praising comments, to encourage academic-task-focussed behaviour and product; and significantly more negative verbal feedback, e.g. critical comments, to discourage undesired social-behaviour ('Off-task' behaviour); and use significantly less positive verbal feedback to encourage students to follow their social-behavioural (non-academic) directions; and significantly less negative verbal feedback to discourage undesired, incorrect or inaccurate academic-task-focussed behaviour or product.

2.72 Research question #2: What other factors are significantly associated with students following teachers' directions ('On-task' behaviour) in academic lessons in UK secondary schools?

2.73 Hypotheses #4, #5, #6, #7, #8, #9, #10, #11, #12 and #13: Under the auspice of Research question #2, these hypotheses are of the same form: 'High levels of time following a teacher's directions in secondary classrooms ('On-task%') are associated with…'; 'Locality type' (hypothesis #4); 'Free school meals %' (#5); 'School roll' (#6); 'Year group' (#7); 'Teacher gender' (#8); 'Years [of teaching] experience' (#9); 'Subject' (#10); 'Adults present [in class]' (#11); 'Class size' (#12); 'AM or PM' (#13). Essentially this is to be an exploratory exercise using both UK1 and UK27 datasets.

2.74 Research question #3: Is the *MICRO* observational tool a valid and reliable means for measuring the association between a teacher's verbal behaviour; and the length of periods of time students spend following a teacher's directions in classrooms in UK secondary schools?**#3:** Is the *MICRO* observational tool a valid and reliable means for measuring the association between a teacher's verbal behaviour; and the length of periods of time students spend following a teacher's directions in classrooms in UK secondary schools?

2.75 Research question #4: Does the length of periods of time students spend following a teacher's directions in classrooms as measured by *MICRO* ('On-task%'); and a quantitative association with teachers' verbal feedback behaviour, and with other non-verbal, contextual factors, have utility for the senior leadership teams in secondary schools?

2.76 Research question #5: Are the length of periods of time students spend following a teacher's directions in classrooms as measured by *MICRO* ('On-task%'); and the positive and encouraging verbal behaviour of teachers reasonable indicators of the conduciveness of the social-emotional climates of classrooms to student learning in UK secondary schools?

2.77 Research question #6: What is the extent to which secondary school teachers might expect their students to follow their directions in academic lessons (as indicated by the 'On-task%' variable) by year group and as a total-mean of the secondary sample?

Chapter 3 – Methodology and method: epistemology, validity and reliability

- **Epistemology and methodological considerations**
- ***MICRO* - Mixed Interval Classroom Observation method**
- ***MICRO* – Validity**
- ***MICRO* - Reliability**
- **Data collection procedures**
- **Iterative Learning Conversations (ILCS) qualitative element using Foucauldian 3x pass thematic discourse-analysis**

3.01 Epistemological and methodological considerations

The decision to use a quantitative method for this study was dependent upon the epistemological standpoint of the researcher, the outcome that was desired and also by the scope-width and accommodative capacity of outcome-generalisations that the researcher wished to make. This research was conducted under the auspices of a rationalist and empiricist epistemology (Nagel, 1961) and positivist ontology (Popper, 1935).

> **3.02** Many of the qualitative methodologies, for example: discourse analysis (Sinclair and Coulthard, 1974; Harris, 1988), thematic analysis (Guest, MacQueen, and Namey, 2012), action research (Elliott, 1991), interpretative phenomenological analysis (Smith, 2009), grounded theory (Strauss and Corbin, 1990), are for convenience limited to small sample sizes and restricted generalisability but can accommodate innumerable or uncounted sets of complexly-interacting, opaquely-defined variables.

3.03 With the research described here, the intent was to create professional impact within the educational domain, and public and political impact by employing a large number of classroom observations using a reductive set of precisely defined variables such as 'On-task%'[16] and four types of verbal feedback used by teachers to indicate the more general discursive characteristics of the social-emotional climate and thus provide a quantifiable measurement of it. Such measurements adhere to MacIntyre's variable limitation caveat (MacIntyre, 1980) considered previously (Chapter 2).

3.04 The method used in this research was quantitative but the statistically-analysed draft results were subsequently examined by a meeting of school managers and the author. The

[16] Variable names that have been designated by the author are displayed in the text within single quotation marks, e.g. 'On-task%' and 'INX%'.

verbal interactions of this meeting were transcribed and thematically analysed (Harding, 2013) using a Foucauldian discourse-analysis model (Foucault, 1969).

3.05 Whilst the generalisability of findings and public impact aspired-to by this research supported the use of a nomothetic quantitative methodology, by adopting MacIntyre's corollary, the results invite criticism of not properly representing a complex idiographic psychological reality: the multivariate social emotional climate for learning in secondary classrooms. It seems unlikely that this criticism can be comprehensively responded-to without recourse to a lengthy discussion of the philosophy of positivist science and methodological reductionism (Popper, 1935; Nagel, 1961). Such a discussion is beyond the practical scope of this chapter. However, it is worth quoting Nagel who provides a philosophical underpinning to MacIntyre's (1980) view and the methodological rationale adopted for the present research:

> However acute our awareness may be of the rich variety of human experience, and however great our concern over the dangers of using the fruits of science to obstruct the development of human individuality, it is not likely that our best interests would be served by stopping objective inquiry into the various conditions determining the existence of human traits and actions, and thus shutting the door to the progressive liberation from illusion that comes from the knowledge achieved by such inquiry. (Nagle, 1961, p.606)

Nagel viewed science as providing a level of explanation that indicates an objective reality but that science could not fully reveal that reality. Nagel's view is less speculative than MacIntyre's position which implies the more variables measured, the more completely a scientist has managed to describe an objective nomothetic reality. Nagel disputes this suggesting that a subset of independent variables will always be beyond the control of the scientist. According to Nagel, science is always contextually limited by time and place to the number of variables that can be observed, measured and controlled for, but if scientific research has been well-designed, it can provide a valid indication of an objective reality, e.g. the social-emotional climate for learning in the classroom, despite findings being limited by constraints of method and limitations upon the number of variables that can be considered by a positivist method.

3.06 The choice of methodology and design of method described here have also been justified by the extended review of research literature in the preceding chapter. This review

provides a detailed historical narrative that logically predicates the current research, its methodology and primarily quantitative method.

3.07 Despite the ancestry revealed in the literature review, large quantitative studies of this type within the domain of educational psychology research appear to be less popular currently (Madill and Gough, 2008), possibly because of the epistemological and ontological standpoints of researchers and current academic fashions thereof, e.g. social-constructionism and constructivism (Berger and Luckmann, 1966; Bruner, 1996) but also because of the practical difficulties of generating large samples and possibly because of the methodological difficulties of isolating any variables to the satisfaction of statistically sophisticated critics and qualitative psychological researchers. Consider, for example, Rutter's influential study: *15,000 Hours* and the theoretical panning that academic statisticians have given the study since publication (Rutter, 1979; Simpson, 1980; Burgess, 1981). The repeated critique of Rutter's method did not diminish the study's continuing persistent and pervasive influence (Wragg, 1999). It is the large scale of such quantitative studies that appears to be an important factor in affording them socio-political impact and influence.

3.08 *MICRO*: Mixed Interval Class Room Observation schedule

3.09 Two time-sampling procedures have been used in developing the observational schedule: *MICRO* (Mixed Interval Class Room Observation), used in the present study.

The time-sampling approach is not without its critics, as has been noted in previously but it is established that momentary time sampling (MTS) is the most reliable variation (Chapter 2).

However, dependent on how the sampling time interval and how it has been determined, the validity of MTS is queried by Ostrov and Hart (2013):

> A major disadvantage of the time sampling approach is that the researcher delineates the particular time interval and therefore arbitrarily categorizes the behavior into discrete artificial units of time that may or may not be meaningful. Moreover some behaviors may exceed the often brief interval of time that is selected for the sampling. Thus, it is crucial to carefully justify the interval selected. (p.287)

3.10 *MICRO*: validity

3.11 The *MICRO* MTS observation of the behaviour of 5 randomly selected students is binarily exclusive – students are judged to be either following the teacher's directions or not: on-task or off-task[17]. There are no partial states.

Assuming a hypothetical class of students managed a low mean 'On-task%' of 66%, an average student in that class would be following directions for 20 minutes of each 30 minute observation period. Assuming that the student was repeatedly on-task for 2 minutes and then off-task for 1 minute, a 1 minute 55 second gap between each 5 second round of MTS observations would mean that the 'On-task%' figure for this hypothetically average student at the end of 30 minutes would be calculated by a hypothetical observer correctly as 66%. Assuming a higher class mean 'On-task%' of 80%, and a student who was repeatedly on-task for 8 minute and off-task for 2 minutes, a 1 minute 55 second gap between each 5 second round of MTS observations would mean that the 'On-task%' figure for this hypothetically average student at the end of 30 minutes would be calculated by a hypothetical observer correctly as 80%. Assuming a 1 minute observational cycle (55 seconds plus 5 seconds) a similar result is obtained. Generating random hypothetical student 'On-task%' patterns shows that the hypothetically observed percentage using MTS remains accurate using a 1 or 2 minute observational cycle and a minimum 30 minute observation period. This provides the rationale requested by Ostrov and Hart (2013).

3.12 In contrast to observing every student in each class (White, 1975; Persons *et al.;* 1976; Merrett and Wheldall, 1986, 1987; and Wheldall, *et al.*, 1989), the author's *MICRO* schedule utilised MTS observations of a randomly-selected sample of 5 students every 2 minutes in each class as an indication of the whole class's behaviour in following a teacher's directions ('On-task%').

The *MICRO* schedule was developed by the author informed by Nafpaktitus, *et al.*'s (1985) large-scale, 84 lesson secondary-phase Californian study; and informed by the author's previous large-scale (141 lessons) primary-phase UK study (Apter *et al.*, 2010) and the observation schedule developed by the author for that study. Nafpaktitus and colleagues' schedule relied upon the MTS observation of 6 randomly selected students per class. The

[17] In the instructions for *MICRO*, on-task behaviour is described as anything that a student does that is in compliance with, or following the teacher's directions or instructions as perceived by an observer following the *MICRO* schedule. Sitting still and listening can therefore be on-task.

author's original classroom observation schedule: the *TaMBiS TSR* ('Teaching and Managing Behaviour in Schools – Time Sample Record'; Apter, 2004) used in the primary study (Apter *et al.*, 2010) relied upon MTS observations of 3 randomly selected students in each class. This was proportionately a small sample: approximately 10% of an average class size of 26.5 (DfE, 2011); but was agreed adequate for this peer-reviewed study.

3.13 The mean class-size of UK secondary schools in 2014 was 20.1 students (DfE, 2014; 2014 was the final year of data collection for the current research). For *MICRO*, the class sample-size of 5 meant that a mean MTS sampling rate of 25% was achieved using the DfE figure as denominator. The DfE does not quote ranges or standard deviations (SD) for mean class-size. The present study can be precise: 236 classes were observed and range was between 4 and 30 students per class; the mean was 19.38; and the SD was 7.07.

This means the representativeness of the random sample of 5 students in each class ranged from 17% to 100% with a mean representation of 26% (see Chapter 4 – Results). Using this *MICRO* MTS sample rate for the study (236 classes of 4574 students, and an observed sample of 1180 students), gives a sampling margin of error (E) of plus or minus 2.45% at a 95% confidence level.

3.14 Observers were asked to observe each lesson for 30 minutes. Adhering to Persons *et al.*'s (1976) innovation, once each MTS observation-cycle of 5 randomly-selected students was completed, the observer switched attention to the teacher's verbal feedback behaviour for the remainder of the two minute time-slice.

The observation of teacher verbal-behaviour was achieved using partial-time interval sampling (PTI) - more accurately, periodic event tallying (PET). During the remainder of each observational cycle, e.g. 1 minute 55 seconds of a 2 minute cycle, the *MICRO* PET was directed at the teacher's verbal behaviour. The observer made 4 separate simultaneous tallies of verbal feedback comments that were praising or critical, academic or social; and a separate simultaneous estimate of the total time talking as a percentage. The 5 tallies were recorded on the *MICRO* sheet:

1. 'TPP'; Task Performance Positive - positively-toned academically-focussed encouragements.
2. 'TPC'; Task Performance Criticism - negatively-toned academically-focussed criticisms.
3. 'SBP'; Social Behaviour Positive - positively-toned comments that recognise and appear to seek to encourage positive social behaviour.

4. ‘SBC’; Social Behaviour Criticism - negatively-toned comments that criticise and appear to seek to discourage negative social behaviour.
5. ‘INX’; INstruction, eXplanation or eXposition: a tally (1 to 5) that is converted to a total talking time percentage, e.g. 3 out of 5 ‘INX%’ equals 60%.

The PET protocol used by *MICRO*, unlike ‘On-task%’ was not binarily exclusive. The four verbal feedback comment-types (‘TPP’, ‘SBP’, ‘TPC’, ‘SBC’) are simple tallies of comments.

3.15 *MICRO*: reliability

3.16 In 2011, prior to the data-collection phase, tests of inter-observer reliability (TIOR) of *MICRO* were completed in two pilot studies on different sites with four observer pairs.

The first TIOR involved 3 pairs of observers using the *MICRO* in Year 7 classes in an average-sized (approximately 900 students) inner-city secondary school in Liverpool. The observers had been trained by reading standardised instructions prepared for distant observers recruited via the internet for the research data-collection phase. The intention was that the TIOR observers would learn and use *MICRO* under the same circumstances as the observers recruited for the project’s data-collection phase. Each pair of observers used the same 5 students for the MTS observations so Cohen’s *kappa* (Cohen, 1960) was an appropriate measure of inter-observer reliability:

Observation pairing	Cohen’s *kappa*	% Agreement
Pair #1	0.79*	91.3 %
Pair #2	0.81*	93.5 %
Pair #3	0.86*	95.6 %

*A kappa above 0.75 indicates a high level of agreement

Table 7 – First test of inter-observer reliability (TIOR) between observations of 5 students’ following a teacher’s directions using *MICRO* to derive ‘On-task%’

Cohen’s *kappa* is not an appropriate measure for the comparison of unstandardised tallies. Observers’ categorisations of teachers’ verbal behaviour using *MICRO* were compared for reliability using a simple percentage agreement formula:

Observation pairing	% Agreement
Pair #1	91.6 %
Pair #2	93.5 %
Pair #3	90.8 %

Table 8 – First test of inter-observer reliability (TIOR) between tallies of different categories of teachers' verbal feedback using *MICRO*

3.17 The second TIOR used two pairs of observations of Year 9 students in a midlands secondary school of approximately 1000 students. Observers were not restricted to observing the same students in each class for the MTS phase of each observational cycle and separately selected students. They did not share their choices with each other. They sat away from each other during observations and secretly filled-in *MICRO* observation sheets. This meant that the TIOR was focussed on the facility of the *MICRO* schedule as an integrated measure. Thus a correlation in each of 2 lessons was sought between the 2 observers' observations in a single calculation that included all *MICRO* categories ('On-task%' behaviour and teacher's verbal-feedback: 'TPP'; 'TPC'; 'SBP'; 'SBC') using a non-parametric test, Spearman's *rho*:

Observation pairing	Spearman's *rho*	% Agreement
Pair #1	0.969*	99.8 %
Pair #2	0.831*	96.7 %

*Significant at 0.01 level (two-tailed)

Table 9 – Second test of inter-observer reliability between two observers using *MICRO* and correlating all observational categories

Whilst Cohen's *kappa* (Cohen, 1960; Bakeman and Gottman, 1997) is first choice for testing inter-rater reliability on single scales, Spearman's *rho* provides a useful test when the scales of categories to be compared were mixed in type, e.g. tally plus ordinal as with *MICRO*; or have wide and/or clustered ranges (Liao, Hunt and Chen, 2010).

3.18 Observer-effects that are inevitable with classroom observation were considered when designing the *MICRO* schedule. Predicating research (Chapter 2) termed this type of observation 'naturalistic' and the aim was to measure the 'natural' rates of 'on-task' behaviour, and frequencies of different types of verbal feedback comments normally used

by teachers in the classroom (White, 1975; Merrett and Wheldall, 1987; Harrop and Swinson, 2000).

Naturalistic observations of classroom behaviour usually involve locating observers in the classroom to be studied. Findings are confounded by participant and observer expectations ('Hawthorne Effect', Mayo, 1933; Roethlisberger and Dickson. 1939; Draper, 2014; and 'Pygmalion' or 'Jastrow Effect', Rosenthal and Jacobsen, 1968).

3.19 Some observational studies of classroom behaviour have experimented with using video cameras and digital audio recording to negate or reduce observer effects (Jewitt, 2012; Swinson and Harrop, 2012), but the large nomothetic scale of the current study, the dependence on the voluntary sign-up of observers and the expense of providing video equipment precluded consideration of this innovation. It was also the aim of the author to develop and validate the *MICRO* schedule as a portable give-away device that psychologists, teachers and academics could use in schools for their own purposes, e.g. teacher-peer observation and mentoring.

3.20 Another argument answering potential methodological 'observer-effect' criticism was that all classrooms observed for the current study had an observer located within them and the same research-purpose described to teachers. Whilst an expectancy effect was expected – given all classrooms observed were probably on 'best behaviour' or at least better than a hypothetical (unobserved and unobservable) 'natural' baseline, this did not preclude being able to compare 'like-with-like' for this study (see Chapter 6 – Discussion).

In conclusion, the inter-observer reliability of the *MICRO* observational schedule is claimed to be valid and reliable for the purposes of the present study.

3.21 Data collection procedures

3.22 Two samples were used in the present study: an observer-selected initial sample of 111 secondary-school academic classes in 27 UK schools, reduced to 106 observations (see Chapter 4) that were statistically valid and representative: the UK27 dataset; and an initial sample of 129 academic classes reduced to 122 academic lesson observations (61 pairs of observations of each teacher and their classes, A.M. and P.M.; see Chapter 5) in one UK midlands-city secondary school selected by the author: the UK1 dataset.

3.23 For the UK27 dataset, psychologist-observers were recruited using the internet discussion forum, *EPNET* (Educational Psychologist Network[18]). Observers identified a familiar secondary mainstream comprehensive school which was neither identified as 'failing' by Ofsted[19] nor identified as especially behaviourally challenging by Local Authority advisers.

Two incentives were offered: access to the UK27 dataset on completion of the research; and entry into a draw for shopping vouchers. This method devised by the author, was used to good effect with the author's primary-phase study (Apter, *et al.*, 2010). The first invitation to take part was posted to *EPNET* in January 2011 and the invitation was repeated approximately every 3 months during the next 3 years.

3.24 Observers were asked to identify a teacher, who they might be able to observe twice teaching an academic subject within Years 7 to 12, once in the morning and once in the afternoon. It was left for observers to choose teacher, year-group and academic subject in the belief that it would be difficult to impose a pre-selection criterion on observers and a prediction that the sample in its entirety would be large enough to ensure a cross-representation of subjects and year groups (Chapter 4).

Only observations of 'teach, talk and write' academic lessons were included in the UK27 and UK1 datasets. Music, P.E., I.T., food technology and art were excluded from both datasets. Lessons that did not require students to talk and write, e.g. exclusively oral / aural modern language lessons, film viewings, etc., were also excluded. This means that lessons observed and used for the study were usually English, maths, science, modern languages, history and geography. There was also an 'Other' category for classroom lessons such as citizenship, guidance and general studies.

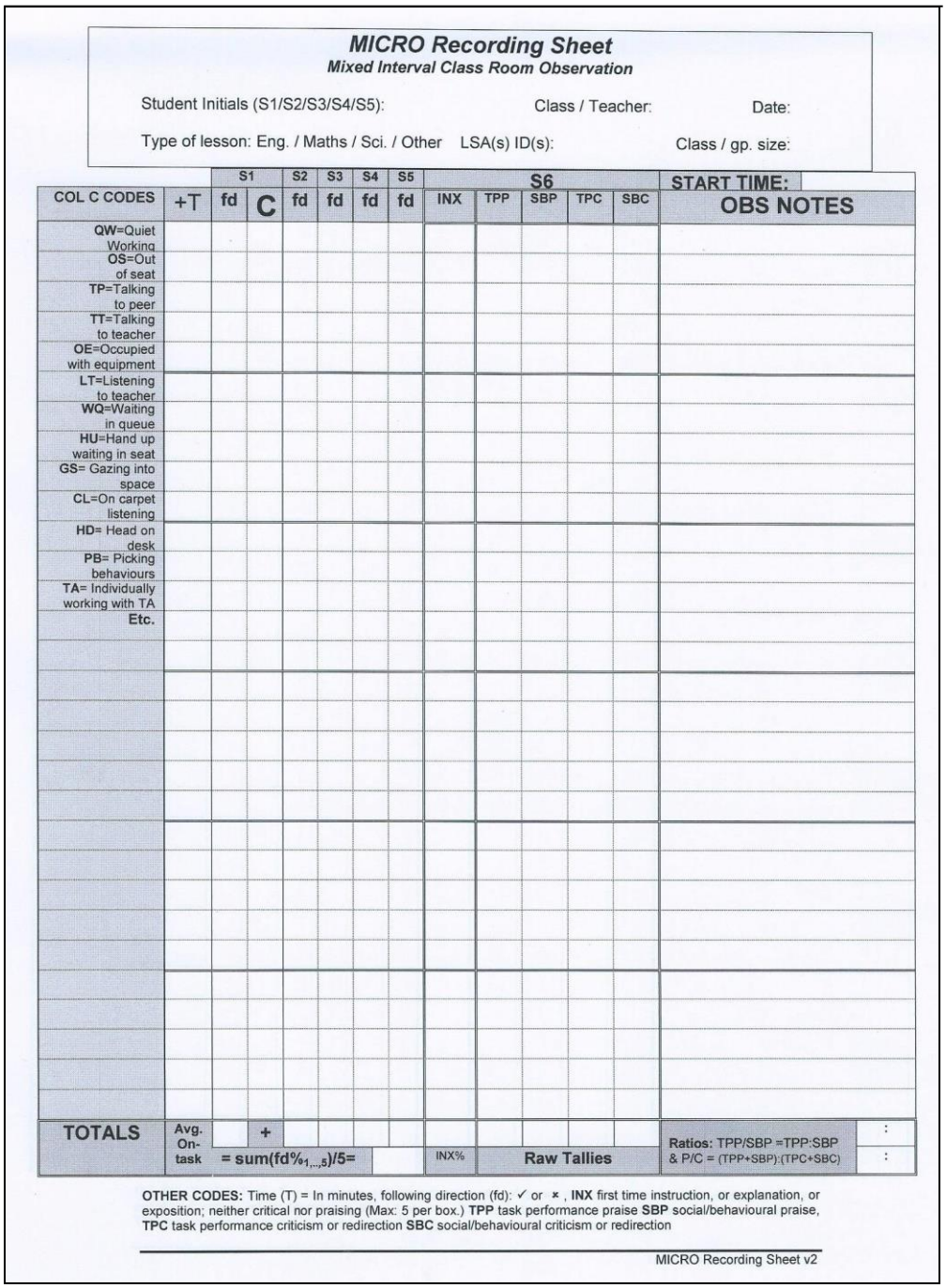
3.25 Observations took place between 2011 and 2014 and 29 observers were who returned at least one pair of observations each admitted to the UK27 dataset. It was hoped that the vagaries of observer-recruitment, choice of school, and classes to observe, meant that dataset UK27 would achieve the same selection-range that would have resulted from a more rigorously defined stratified random sampling method. It can only be accurately claimed

[18] EPNET@JISCMAIL.AC.UK - a professional list / discussion website managed, monitored and maintained by the University of Newcastle.
[19] Ofsted: Office for Standards in Education, Children Services and Skills – the UK state inspection and regulation service

however that an un-stratified quasi-random sample (Bajpal, 1971) of UK secondary academic classes was achieved - years 7, 8 and 9 were marginally over-represented in the UK27 dataset (Figure 20).

3.26 It was important that the *MICRO* schedule could be easily and reliably taught to Educational Psychologists using a manualised training delivery. Educational Psychologists were assumed to have a prerequisite of classroom and student observation skills as part of their professional qualification and experience. The training materials were field-tested and revised a number of times by colleagues in the author's educational psychology service. An email helpline was set up so that observers could send queries to the author. This service was infrequently used by observers throughout the data-collection period.

3.27 The *MICRO* schedule primarily consisted of a observation record that was printed-off as an A4 single sheet (Figure 1). This was accompanied by a manual with exemplars of each teacher comment-category ('TPP'; 'SBP'; 'TPC'; 'SBC') and of neutrally toned teacher-talk ('INX'). Other documents were: an anonymised data-summary sheet; and the template for a permissions-request letter intended for Head Teacher. Purposes, ethical considerations and guarantees of anonymity were explained in the letter. Copies of the 4 documents are appended (Appendix).



MICRO Recording Sheet
Mixed Interval Class Room Observation

Student Initials (S1/S2/S3/S4/S5): Class / Teacher: Date:

Type of lesson: Eng. / Maths / Sci. / Other LSA(s) ID(s): Class / gp. size:

		S1		S2	S3	S4	S5	S6					START TIME:
COL C CODES	+T	fd	C	fd	fd	fd	fd	INX	TPP	SBP	TPC	SBC	OBS NOTES
QW=Quiet Working													
OS=Out of seat													
TP=Talking to peer													
TT=Talking to teacher													
OE=Occupied with equipment													
LT=Listening to teacher													
WQ=Waiting in queue													
HU=Hand up waiting in seat													
GS= Gazing into space													
CL=On carpet listening													
HD= Head on desk													
PB= Picking behaviours													
TA= Individually working with TA													
Etc.													
TOTALS	Avg. On-task		+										Ratios: TPP/SBP =TPP:SBP :
		= sum(fd%$_{1,..,5}$)/5=						INX%	Raw Tallies				& P/C = (TPP+SBP):(TPC+SBC) :

OTHER CODES: Time (T) = In minutes, following direction (fd): ✓ or ×, **INX** first time instruction, or explanation, or exposition; neither critical nor praising (Max: 5 per box.) **TPP** task performance praise **SBP** social/behavioural praise, **TPC** task performance criticism or redirection **SBC** social/behavioural criticism or redirection

MICRO Recording Sheet v2

Figure 1 – *MICRO* recording sheet

3.28 For the single-school dataset UK1, 4 observers were appointed by the author: the author and 3 colleagues.

The UK1 school was familiar to the author. There had been a long positive relationship between school and Educational Psychology Service. The involvement of the school in the current research came about when the Head Teacher and Senior Leadership Team (SLT) requested a survey of how teacher feedback was being used in a service-planning

consultation after receiving an inspection judgement by Ofsted that the school required 'special measures' (ATL, 2013)[20]. Whilst observations of all lessons, year groups and teachers were undertaken, only paired observations (same teacher; A.M. and P.M.) of academic classes were used for the UK1 dataset. Observations were completed, January to May, 2014.

3.29 Iterative Learning Conversations (ILCS) qualitative element using Foucauldian 3x pass thematic discourse-analysis

3.30 A regular senior leadership team (SLT) meeting in the UK1 school was used to generate an Iterative Learning Conversation (Apter, 2014) as a peer-examination of the initial findings from the UK27 and UK1 datasets (Chapters 4 and 5). The author assumed the role of meeting facilitator. The meeting lasted 1.5 hours and began with a presentation by the author of a draft analysis of the UK1 and UK27 datasets. The ILCS transcription of the verbal interactions of the meeting was used as the dataset for a structured examination of how the SLT regarded the emergent quantitative findings of the study. This exercise enhanced the rationalist examination (Nagel, 1961) of the findings of this study that have been reported here (Chapters 5 and 6).

3.31 When individuals were presented with opposing opinions or directly challenged, the facilitator encouraged them to modify or defend their position. In a focus group (Turner, 2012) the meeting facilitator is required to be impartial to the extent that their amplification of issues raised by the group might depress emergent information that was new or previously hidden. In this ILCS meeting (Apter, 2014), the facilitator needed to ensure that opportunities to jointly develop an analysis during the meeting were not under-used as the learning conversation developed and thematically iterated, and the forcefulness of contributions were felt by the group. There is a risk of a consensus being assumed when not every person has spoken or has spoken infrequently. The facilitator needed to consider carefully whether people who had not expressed a view could be assumed to be tacitly agreeing with the majority, or whether they may simply have been unwilling to voice their disagreement.

[20] In response to an agreement with the Head Teacher, the school's identity has been anonymised and the Ofsted inspection report has not been referenced because of the sensitivity of these matters in the light of the findings reported here (Chapter 5 and 6).

3.32 A verbatim transcription of the meeting was made and an ILCS Foucauldian 3x pass thematic discourse-analysis (Apter, 2014) of the transcription completed. With a focus-group approach (Harding, 2013) the analysis of transcripts is commonly an intentionally imprecise thematic examination, but using ILCS, the Foucauldian thematic discourse-analysis applied three successive coding processes to the transcription.

3.33 The first analytical pass through the transcribed text was used to analyse and annotate each line using Iterative Learning Conversations codes (ILCS; Apter, 2014). There are 10 descriptors used by an ILCS analysis to broadly describe each utterance in terms of the type of information that it was carrying in terms of change-activity (Apter, 2014):

Iterative Learning Conversations ILCS categorical codes		
Primary	**A1: Agents / audiences**	Lists of people involved in the change-process[a] – agentically (agents) and non-agentically (audiences – stake-holders and bystanders) in talking/communicating/behaving within the transcribed/annotated ILCS. Agents and audiences are described in respect of identities *and* in respect of their context – the historical, cultural and political "givens" that are assumed by the perceiver/recorder[b] to apply to them.
	A2: Attributions / agency	The agency (evident intentions) of agents; and the assumed knowledge expressed by both agents and audiences – including speech-*acts* (**ACT**, **Iy** or **Py**), represented usually as verbatim adjacency pairs and perceptions from the standpoint of the perceiver/recorder (usually the psychologist) of others' attributions. "Attribution" is used in the Heider (1958) *Attribution Theory* sense of contextualised meanings.
Subsidiary	**P1: Possibilities**	Possibilities of change (problems) that are special types of attributions identified by each agent, including possibilities/problems that are jointly identified by agents or audience members.
	P2: Propositions	These are tentatively proposed functional relationships – usually between possibilities of change (P1) and the elements that make up a context. Differing from the concept of a hypothesis, a *proposition* is a sentence comprised of words that attribute a causal relationship existing between different pieces of information, for example, *the boy feels sad because his mother has died*, or *if we put a plaster on the cut, it will stop bleeding*. In a given ILCS account, there might be multiple related propositions that are deemed worthy of testing.
	T1: Tests	Quasi-objective tests, checks and trials that produce ostensible evidence (E1) that can be associated with propositions (P1) in order to expose their utility. Tests might include behavioural experiments (trying a new behaviour out), data from standardised assessments including questionnaires, graded tasks or exposure to new contexts (for example, trial placements), and verbal checking, for example, *how is that so?*
	T2: Timescales	Historical strings of related events, *timed* and *dated*, including when events, contexts and behaviours have occurred, when assessments and interventions were completed, and when they are planned to occur in the future.
	E1: Evidence	Quasi-objective information and measures – including T1, observations, test and experiment results; and emergent information from trial placements and changed contexts. Evidence should include judgements by the recorder of ecological and systemic validity from *their standpoint*.
	E2: Estimates	Quasi-objective information that is not known has been estimated, projected or regressed from E1 and the identification of information that is still required. E2 is usually attributed information and thus can be expressed as a factor of A2, thus: A2(E2).
	R1: Reflections	Emotional, social and intellectual *post hoc* reflections of agents and audiences including the reflections of the ILCS recorder (usually the psychologist) which might be informed by cited research – expressed optionally as A2(R1).
	R2: Responses	A punctuation at a hypothetical endpoint of an ILCS, as it has been signified by the recorder; including planned next steps, reboots of ILCS, repeats of sections (re-iterations) and the reasons for so doing attributed by agents/audiences. A reiteration might branch back to an earlier specified point in the ILCS. A response at the end of an ILCS might include a planned review cycle or case-closure.

[a]Every socio-historical process is a *change* process, including organisational management, personal therapy, and governance.
[b]A perceiver/recorder is an ILCS practitioner who might also be a psychologist, management consultant, therapist or politician.

Figure 2 – ILCS Categories (Apter, 2014, p. 338)

3.34 The second pass employed a variation of Foucauldian discourse-analysis (Foucault, 1969; Apter, 2014) using Speech Act codes for locutionary, illocutionary and perlocutionary elements (Austin, 1962). This second-pass analysis was used to reveal and describe emergent power-plays and the use of verbal forcefulness within the discourse.

Austin (1962) proposed that verbal behaviour was comprised of a series of performatives intended to have an effect on others. In order to annotate how power is utilised in conversations, ILCS uses Austin's Speech Act Theory to categorise performatives into three integral elements called speech acts:

1. The locutionary act - **ACT**: the performance of an utterance: the actual utterance and its ostensible meaning, comprising the verbal, syntactic and semantic aspects of any meaningful utterance.
2. An illocutionary act - **Iy**: the pragmatic force of the utterance, in respect of its intended significance as a socially-valid verbal action.
3. A perlocutionary act - **Py**: its intentional effect, for example, persuading, convincing, scaring, enlightening, inspiring, or otherwise getting someone to do or realise something, whether intended or not. The perlocutionary act is implicit in the illocutionary act but is gauged once a behavioural response, usually a verbal response, has been made by a respondent.

3.35 The third pass employed colour-codes (Harding, 2013) for emergent themes pragmatically identified by the author as having particular relevance for the hypothesis and research questions of the current research (Chapter 2).

The 10 themes and subsidiary illustrative thematic questions are tabulated (Table 29; Chapter 5).

3.36 The transcription of what was said in that meeting with annotated ILCS 3-pass Foucauldian codings is appended (Appendix).

The summative conclusions of the analysis are presented after the quantitative analysis of the comparison between the UK27 and UK1 datasets (Chapter 5).

Chapter 4 – UK27 Results - analysis of classroom observations in 27 UK secondary schools (UK27 dataset)

- **UK27 dataset descriptive statistics**
- **UK27 dependent variable ('On-task%')**
- **Graphs of UK27 teacher verbal feedback variables**
- **Hierarchical multiple regression analysis (HMRA) of UK27 teacher-feedback variables and 'On-task%'**
- **UK27 teacher-feedback and 'On-task%' correlations**
- **UK27 teacher-feedback type ratios**
- **UK27 other teacher-feedback correlations**
- **UK27 contextual variables descriptives**
- **UK27 HMRA of contextual variables**
- **Other correlations in the UK27 dataset involving contextual variables**
- **UK27 teacher-feedback by year group**
- **UK27 teacher-feedback by subject**

4.01 UK27 dataset descriptive statistics

The UK27 dataset was compiled from *MICRO* data-summary sheets returned to the author by 29 observers of 111 academic lessons in 27 secondary schools in a range of locations including Plymouth, London, Birmingham, Liverpool, Glasgow and Cardiff.

One set of observations was excluded from the UK27 dataset because a researcher returned data-summary sheets that indicated that they had misread the instructions that they had been sent and appeared to have reversed one or more scales used by the *MICRO*. One other set of observations was excluded from UK27 because data was incomplete.

One observer's data from one of their pair of observations was subsequently removed from UK27 for the statistical procedures reported in this chapter because the Mahalanobis distance calculated for their data exceeded the critical value for 5 variables used for teacher verbal feedback analysis and their effects upon the dependent variable ('On-task%'). These data were not subsequently re-included for other analyses included in the current research; for example: the analysis of the effects of 10 contextual variables upon the dependent variable, even though the calculated Mahalanobis distance did not exceed the critical value for 10 variables (Tabachnick and Fidell, 2012).

4.02 The following 16 variables (5 teacher verbal feedback variables, 10 contextual variables, and the singular dependent variable 'On-task%') were included in the analyses of 106 classroom observations:

	Variable Name	*Description*	*Range*
d.v.*	On-task%	Following teacher-directions mean-time as %	0 - 100
Teacher verbal behaviour – independent variables	TPP	Task performance positive teacher-comment	0 – 49+
	SBP	Social behaviour positive teacher-comment	0 – 15+
	TPC	Task performance criticism teacher-comment	0 – 25+
	SBC	Social behavioural criticism teacher-comment	0 – 39+
	INX%	Neutral teacher-talk as a proportion of time %	0 – 100
School and context - independent variables	Locality type	Inner city, suburban or rural	1, 2 or 3
	Free school meals%	Proportion of school roll %	0 - 100
	School roll	No of students on school roll	543 – 1683+
	Year group	National curriculum year of class	7 - 12
	Teacher gender	Male / female	1= m. or 2= f.
	Years' experience	Number of years teacher has taught	Less than 1 to 25
	Subject	English, maths, science etc.	1,2,3,4,5,9,10*
	Adults present	Number of adults in classroom	1 - 6
	Class size	Number of students in classroom	4 – 30+
	AM or PM	Lesson in morning or afternoon	1= a.m. 2= p.m.

*UK27 Additional Codes: d.v.= dependent variable, m.= male, f.=female, 1= English, 2= Maths, 3= Science, 4= History, 5= Geography, 6= Modern language, 7= Music 8= Other - a catch-all category that includes more informal lessons such as pastoral guidance and citizenship that do not culminate in an examination

Table 10 - Key to UK27 data variable types

4.03 The following table displays the data limitations of the UK27 dataset in terms of simple descriptive statistics and minimum / maximum recorded values[21] for 16 variables:

[21] IBM Statistical Package for the Social Sciences (SPSS) Version 23 was used for all data-processing. All results from calculations are given to 2 decimal places for primary data and means or 3 decimal places by author's discretion for second order SPSS calculations where the product is small, e.g. standard deviations < 5, etc.

	Variable	N	Min.	Max.	Mean	SD
d.v.	On-task%	106	39.00	99.00	78.15	13.31
Teacher verbal behaviour – independent variables	TPP	106	0	49	16.78	11.11
	SBP	106	0	15	2.89	3.725
	TPC	106	0	25	5.11	6.02
	SBC	106	0	39	12.21	9.87
	INX%	106	15	100	52.66	18.91
School and context independent variables	Locality type	106	1	2	1.43	.497
	Free school meals%	104	5	71	20.86	16.61
	School roll	106	543	1683	955	346.86
	Year group	106	7	12	8.66	1.498
	Teacher gender	106	1	2	1.52	.50
	Years experience	106	0	25	8.20	5.87
	Subject	106	1	10	4.33	3.21
	Adults present	106	1	6	1.33	.75
	Class size	106	4	30	19.43	7.085
	AM or PM	106	1	2	1.31	.465

Table 11 – UK 27 dataset, descriptive statistics

There were no rurally located schools in the sample. All schools were in suburban, metropolitan or inner-city locations. Schools, classes and teachers were selected by observers as matching the research inclusion criteria provided in advance documentation (appended: Appendix; see also Chapter 3). The mean duration of lesson observations used for compiling the UK dataset was 32.5 minutes (SD= 2.92).

4.04 The following figure is a graph depicting boxplots[22] of 15 variables (N=106). It excludes 'School roll'. The range of school rolls - from 543 to 1683 students - meant that a boxplot could not be meaningfully incorporated in the same graph, but the UK27 sample-mean of 959 and the range of 543 to 1643 provides a good analogue for the Department for Education's (DfE) figures for the UK as a whole with a secondary school roll mean of 955.65 students (DfE, 2014).

[22] Each plotted box represents 50% of cases, with whiskers extending to smallest and largest values. The circles and numbers represent outliers, with values that are more than 1.5x more than the adjacent edge-value of a box. The dark line across the rectangle represents the mean.

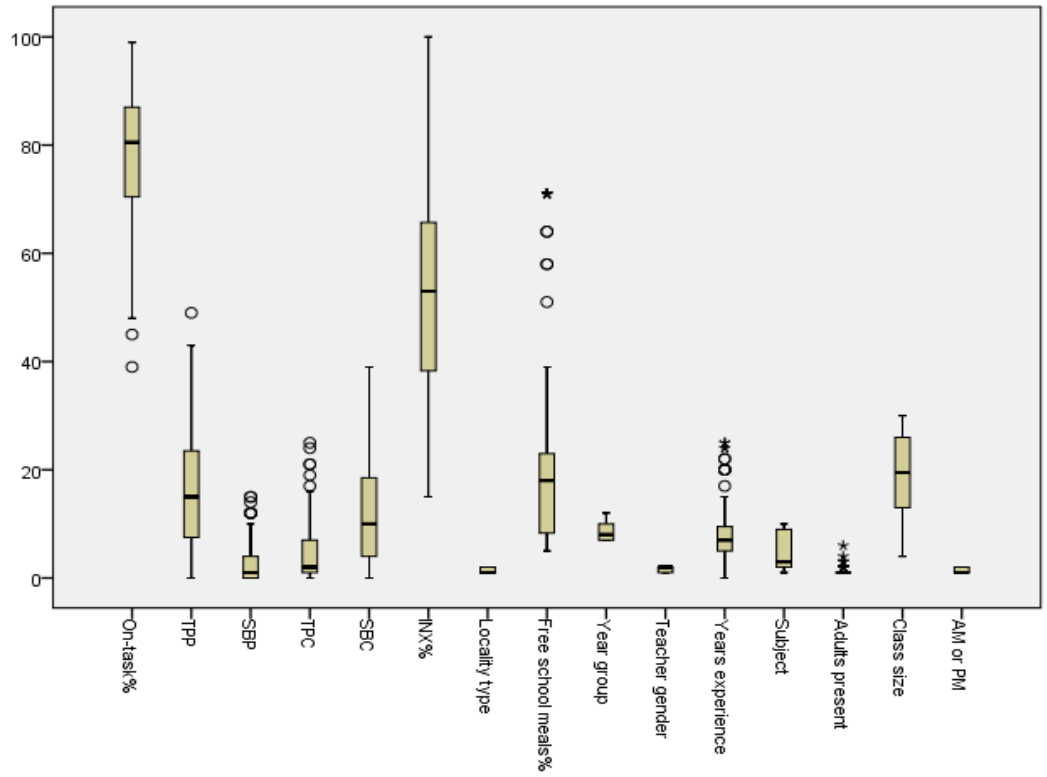


Figure 3 – Boxplots of UK27 variables included in the current research (N=106)

This graph serves to graphically represent data distribution centred on means, and any outliers. There were 2 lower outliers from the dependent variable boxplot, 'On-task%'. With teacher-feedback variables: there were 5 upper-outliers from the boxplot of 'SBP' and 7 from 'TPC'; and with school context variables: 7 from 'Years' experience', 8 from 'Free school meals %', and 4 from 'Adults present'. Whilst calculated as anomalous (as determined by SPSS), independent contextual and feedback variables are indisputable[23] tallies and no other outlier data points were removed.

In the case of 'TPC' and 'SBP' an exponential distribution and upward-outliers was expected from the review of previous literature: teachers have not previously been observed to make significant usage of these feedback-types (Chapter 2).

The smallest boxes represent the 3 dichotomous school context variables: 'Locality type' (that there were no *rural* secondary schools – a category on the data summary form - was an unanticipated 'given'); 'Teacher gender'; and 'AM or PM'.

[23] 'Indisputable' insofar as there are no practical means of disputing these tallies.

4.05 UK27 dependent variable ('On-task%')

The distribution of the dependent variable: 'On-task%' is revealed in the following bar graph:

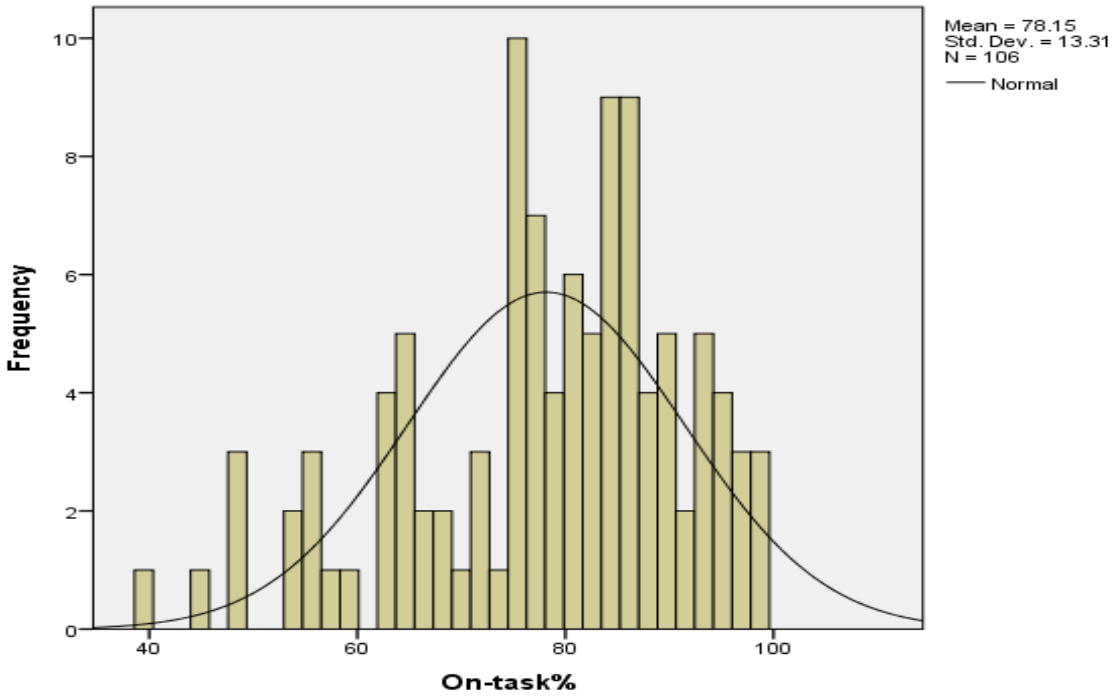


Figure 4 – Distribution of 'On-task%' in UK27 dataset with 'best-fit' normal distribution curve (N= 106)

Testing the UK27 'On-task%' data for being normally distributed, the distribution failed the Kolmogorov-Smirnov test with D (106) = .12, $p < .005$. This might appear to mean that parametric statistical analytics should not be used with the UK27 dataset in the first instance. However, there is a balance to be drawn between the likelihood of a Type 1 error and the potential power of a statistical test (Pett, 2016). Insofar as establishing a predictive relationship between the distribution(s) of one or more independent variables and the distribution of a dependent variable to support a hypothesis, parametric tests are more powerful than their non-parametric equivalents (ibid.).

The 5% trimmed mean of the dependent variable ('On-task%) was 78.83% which was satisfactory (ibid.). Skewness was measured at S= -.76 (SE= .235) which described the midpoint of the distribution being skewed to the right as is evident from the histogram

above (Figure 4) Kurtosis, measured at *K*= .13 (SE= .465) suggested that the distribution was steeply peaked.

4.06 The Q-Q Plot strongly suggested normality whilst reflecting the previously noted caveat about skewness:

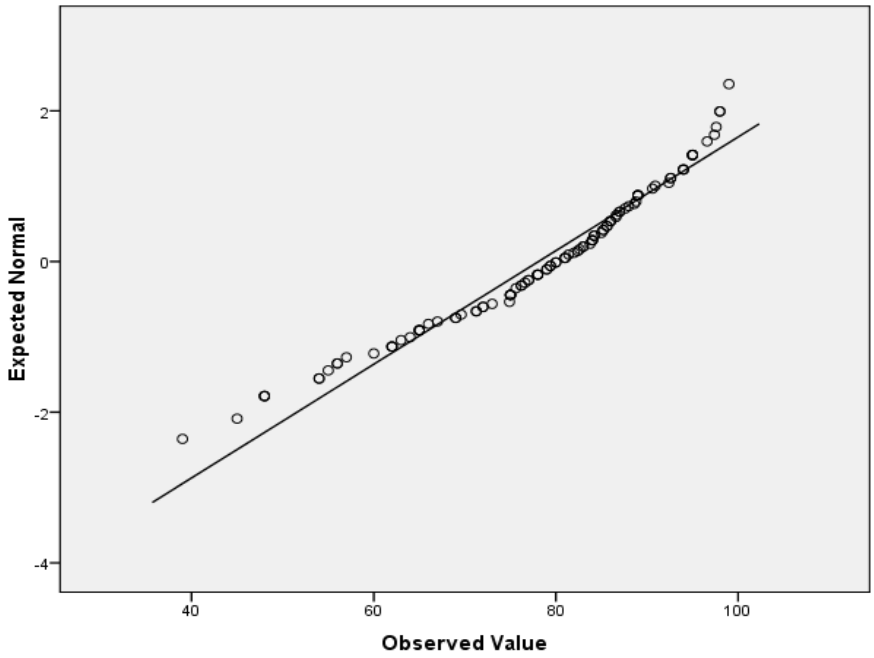


Figure 5 – Normal Q-Q Plot of UK27 'On-task%' variable

Whilst the distribution of the UK27 dependent 'On-task%' variable failed the Kolmogorov-Smirnov test of normality, a number of characteristics supported continued consideration of using the more powerful parametric group of tests (Pett, 2016). These characteristics were: the sample was moderately large (N= 106); the mean was more valid than the median as a significant characteristic of the distribution; and the number of outliers was small (N= 2), not too distant from the median and in the direction of the more complete leading tail - given the right-handed skew of the distribution[24]. With large datasets, smooth distributions, either normal or exponential, are improbable (Field, 2009).

4.07 Notably, the mean following teacher-directions ('On-task%') percentage at 78.15% for UK27 secondary schools including all year groups (7 to 12) was 6.34% lower than the mean

[24] Data was also subjected to non-parametric tests particularly when a distribution severely violated parametric assumptions but where the variable concerned was particularly pertinent to the research questions, e.g. 'TPC' and 'SBP'. Results of these analyses are reported in the main body of the report. Other non-parametric results are reported in the Appendix for the sake of completeness.

'on-task' percentage of 84.49% obtained in the author's similarly large-scale 2010 primary school study (Apter, *et al.*, 2010; N= 141).

4.08 Graphs of UK27 teacher verbal feedback variables

The 8 graphs below are histograms and respective Q-Q Plots representing 4 types of teacher-feedback comments included in the UK27 dataset: 'TPP' (Task Performance Positive); 'SBC' (Social Behaviour Criticism); 'TPC' (Task Performance Criticism) 'SBP' (Social Behaviour Positive):

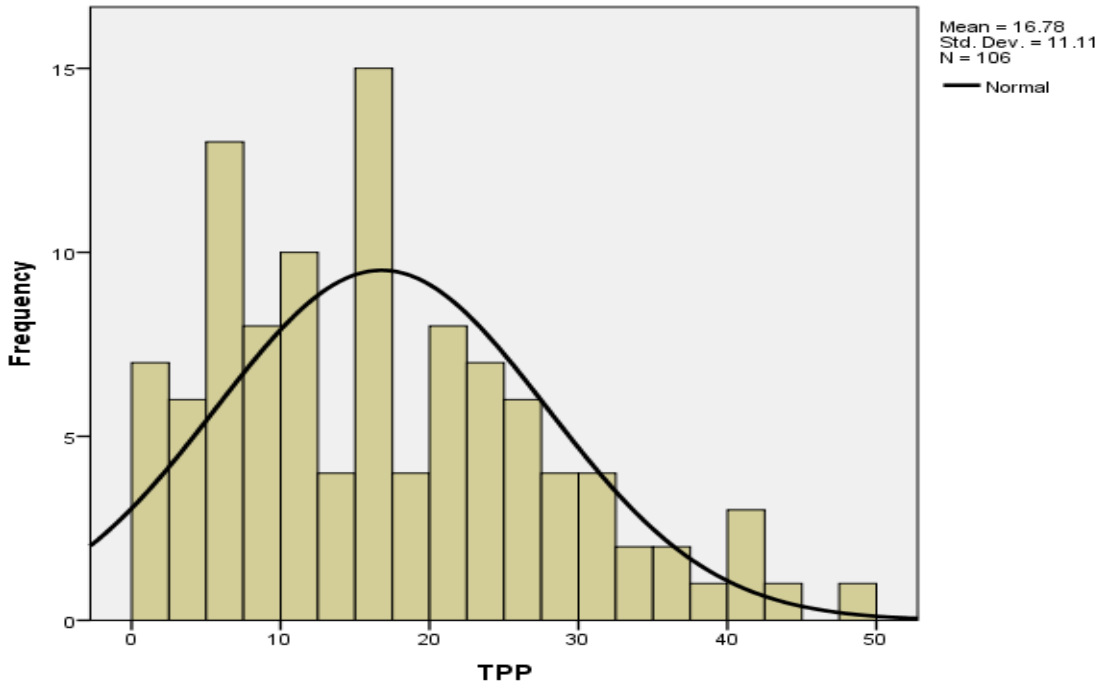


Figure 6 - Histogram representing UK27 'TPP' (Task-Performance Positive) teacher-feedback comments with best-fit normal distribution curve (N=106)

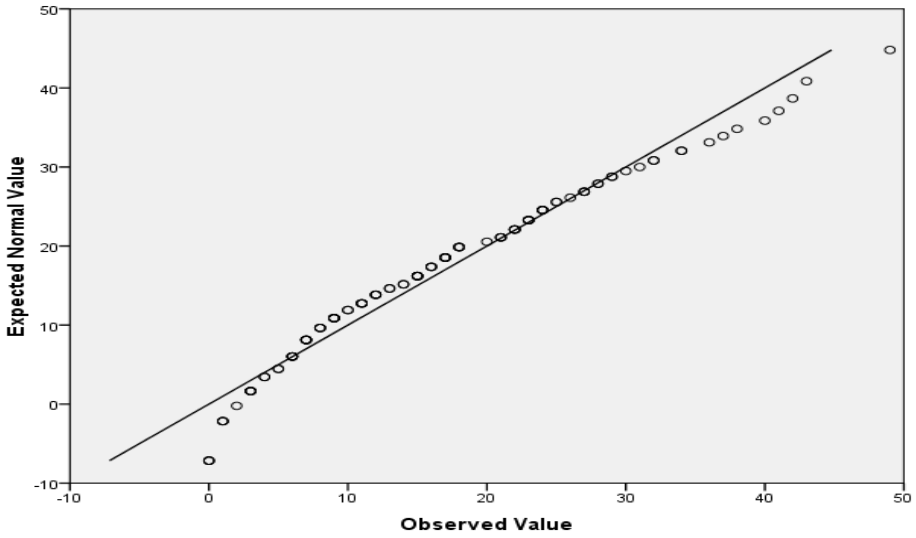


Figure 7 – Normal Q-Q Plot of 'TPP'

The mean and distribution suggests that teachers used positive feedback-comments directed at students task performance ('TPP') at a high enough rate (mean= 16.78, N=106) and with a normal enough distribution to warrant investigating further using parametric methods the association between 'TPP' and 'On-task%'.

4.09 The exponential distribution of teachers' use of social behaviour positive comments 'SBP' and the relatively small *total* number of 'SBP' comments in the UK27 dataset (mean= 2.89, N= 106) reflected the repeated theme in the literature (Wheldall *et al.*, 1989; see Chapter 2) that teachers in secondary schools do not usually congratulate students for good behaviour.

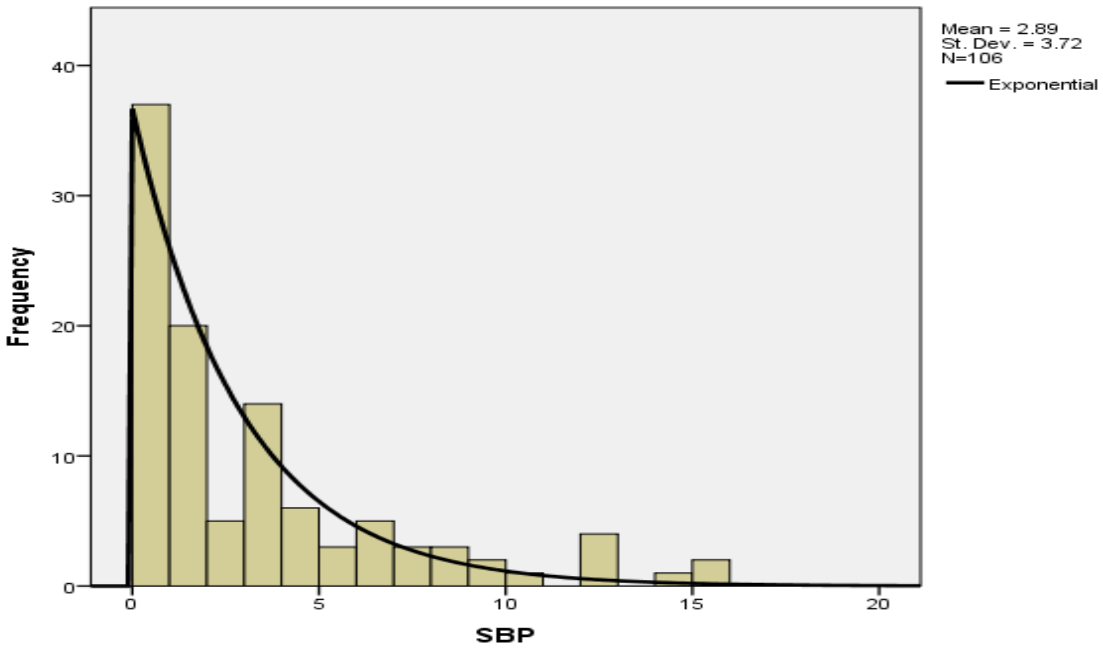


Figure 8 – Histogram representing UK27 'SBP' (Social Behavioural Positive) teacher comments with 'best-fit' exponential distribution curve (N=106)

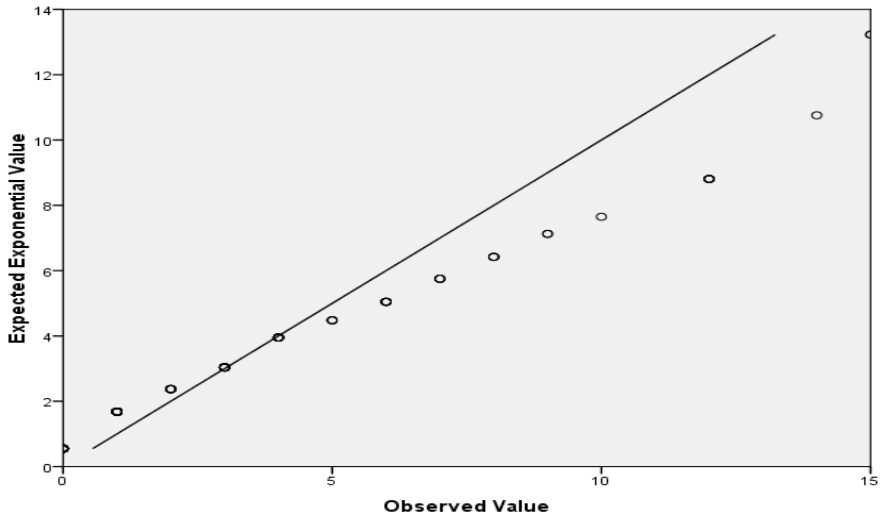


Figure 9 – Exponential Q-Q Plot of 'SBP'

The exponential distribution of 'SBP' warrants further investigation but indicates that a zero instances score was the most frequently occurring case (mode).

4.10 As with 'SBP', the exponential distribution of teachers' use of task performance criticism 'TPC' comments and the relatively small total number of 'TPC' comments in the UK27 dataset (mean= 5.11, N= 106) reflected another repeated theme in the literature (Wheldall *et al.*, 1989; and see Chapter 2) that teachers in secondary schools rarely criticise

students for their academic work – at least in respect of making comments that the whole class and an observer would hear.

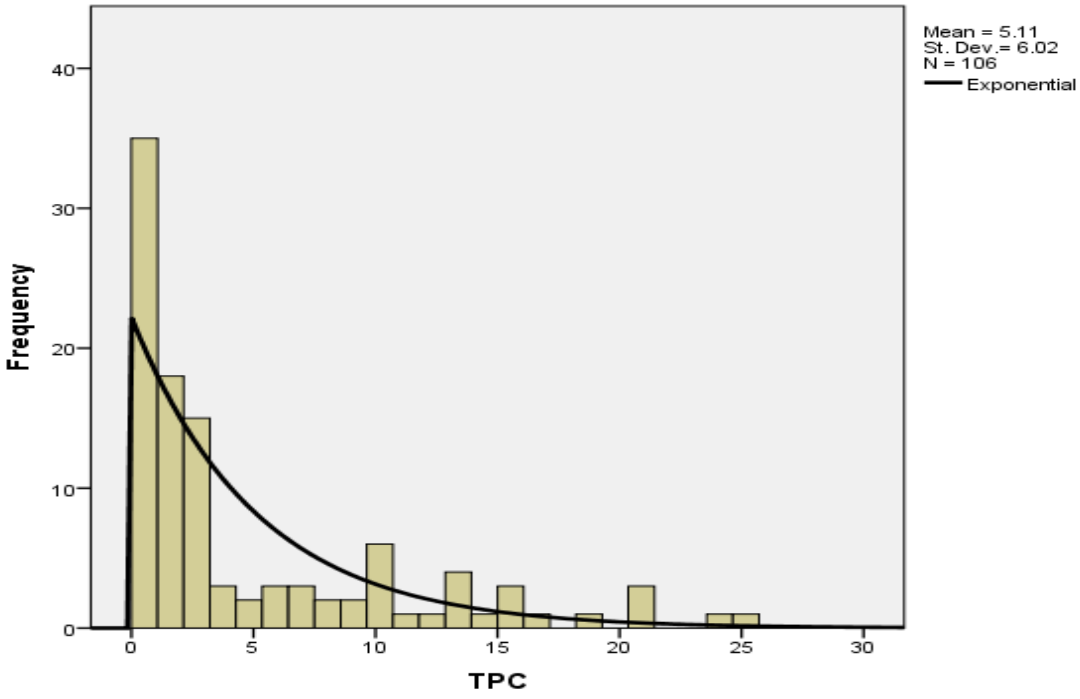


Figure 10 – Histogram representing UK 27 'TPC' (Task Performance Criticism) teacher comments with 'best-fit' exponential distribution curve (N= 106)

Once again, the exponential distribution of 'TPC' warrants further investigation but indicates that a zero-instances score is the most frequently occurring case (mode).

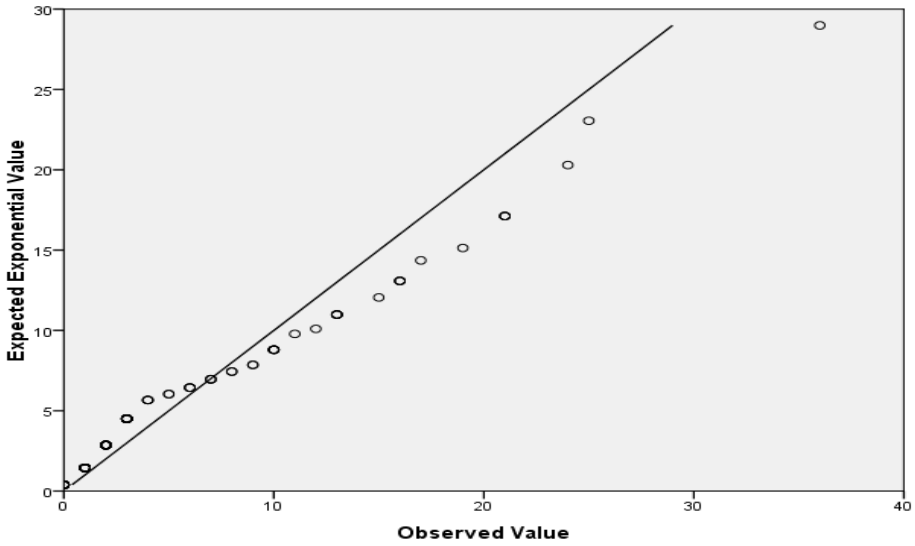


Figure 11 – Exponential Q-Q Plot of 'TPC'

4.11 With ‘SBC’ – teacher-feedback comments that were critical of students social behaviour, higher frequencies of comments and a lower proportion of zero scores is represented in the following histogram:

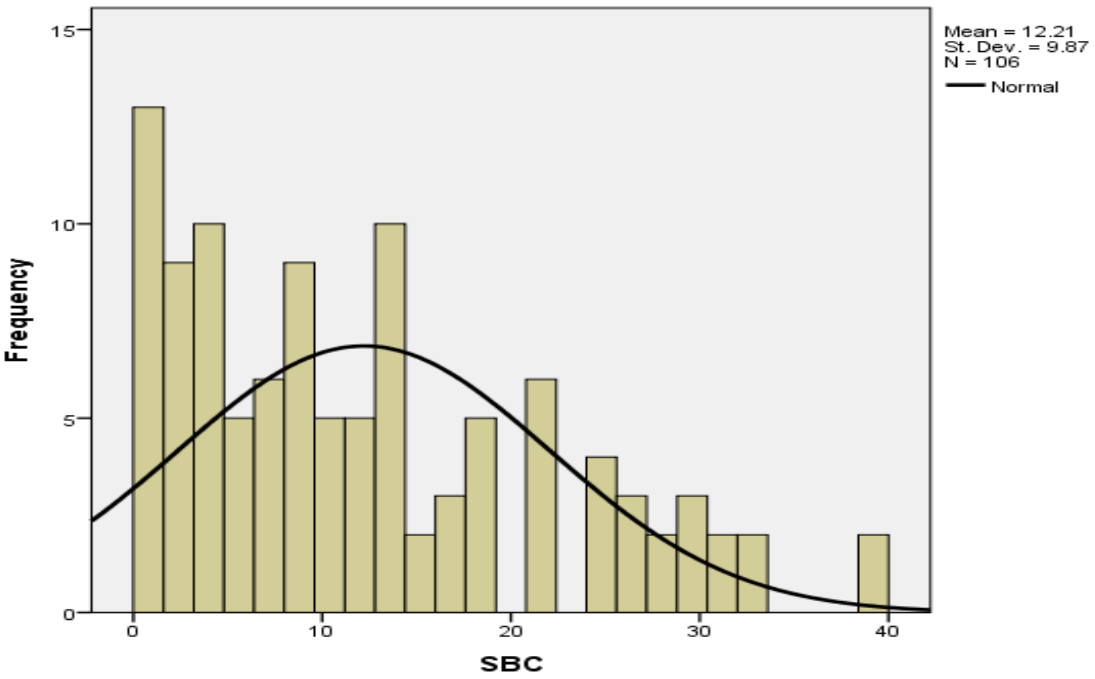


Figure 12 – Histogram representing UK27 ‘SBC’ (Social Behaviour Criticism) teacher comments with ‘best-fit’ normal distribution curve (N=106)

The UK27 ‘SBC’ mean at 12.21 (SD= 9.87; N= 106) appears much higher than the means for 2 teacher-feedback tallies ‘SBP’ (mean= 2.89); and ‘TPC’ (mean= 5.11) but lower than ‘TPP’ (16.78). These differences are analysed as ratios and for significance below.

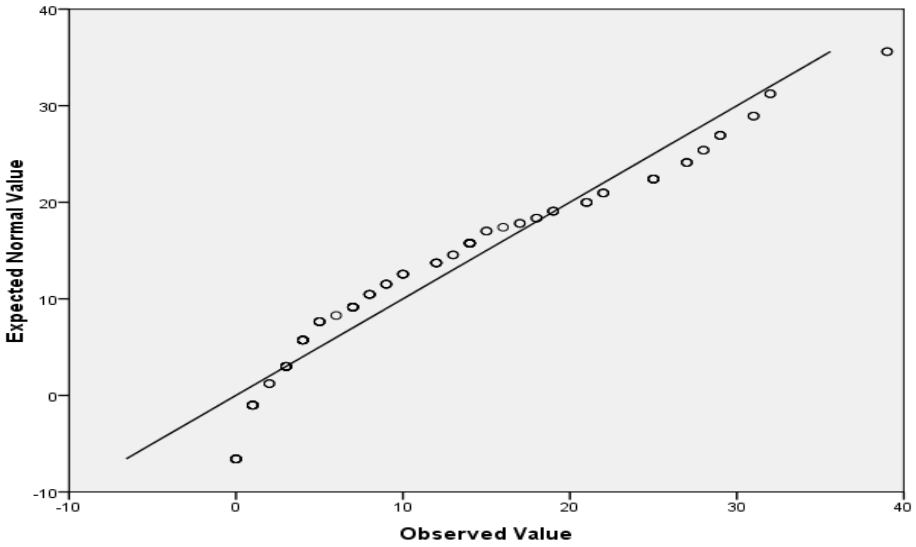


Figure 13 – Normal Q-Q Plot of UK27 'SBC'

4.12 An estimation of the mean amount that teachers spoke for in the UK27 dataset lesson-observations - irrespective of their use of positive or negative feedback comments to students, was captured in the 'INX%' variable (neutral teacher-talk as a proportion of time %). The histogram and Q-Q Plot below depict a steep normal characteristic in the distribution:

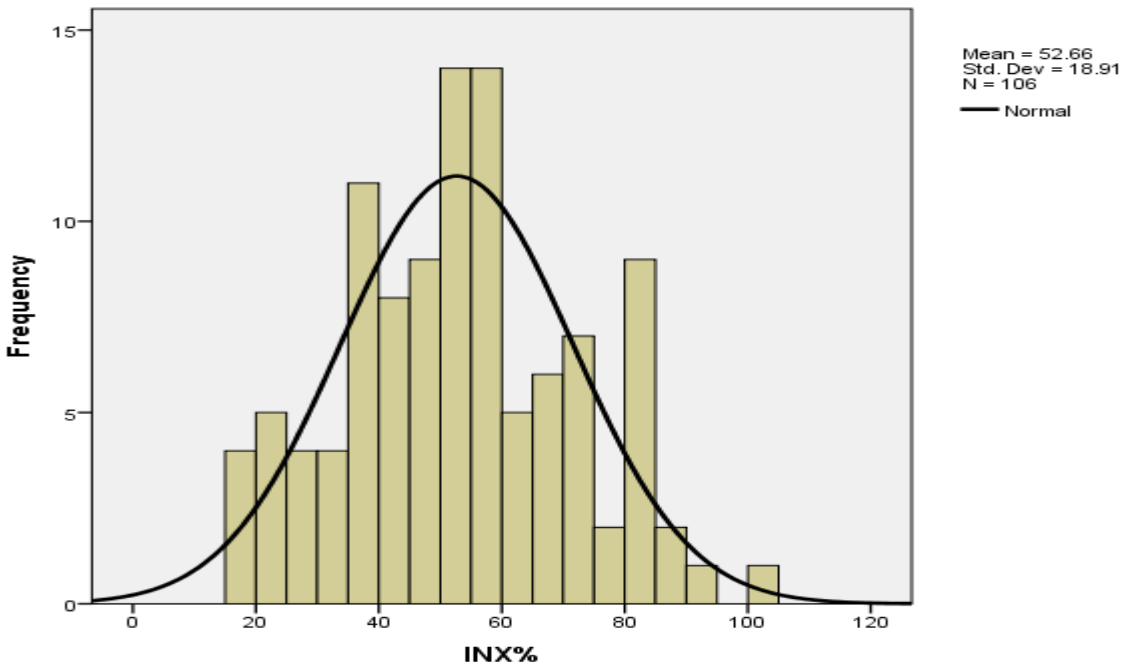


Figure 14 - Histogram representing UK27 'INX%' neutral teacher-talk as a proportion of time % (N=106)

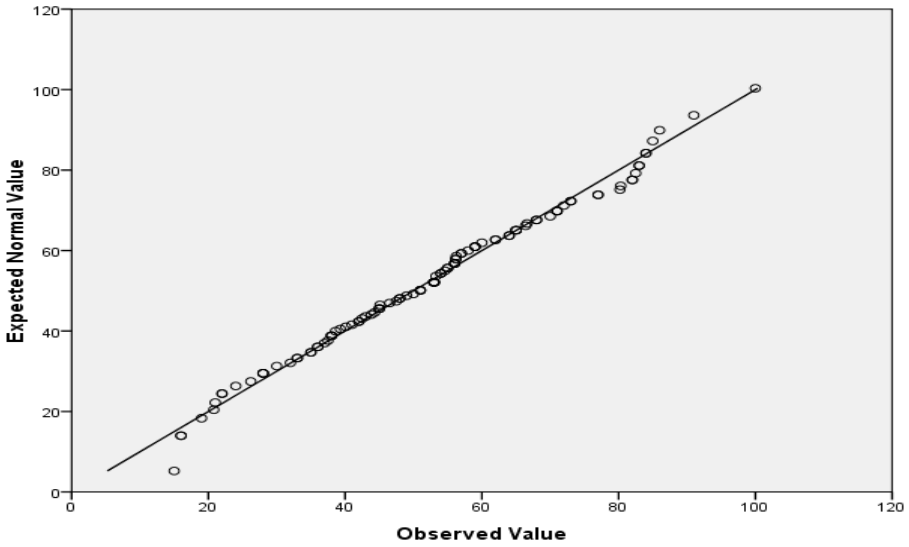


Figure 15 – Normal Q-Q Plot of UK27 'INX%'

4.13 Whilst some of the verbal feedback data, e.g. particularly 'SBP' and 'TPC', technically violated the parametric requirement for normality, the parametric statistical method of choice to identify the predictive relationship between a number of different dynamically independent variables, e.g. the amount of neutral teacher-talk: 'INX%'; and teacher-feedback comments: 'TPP'; 'SBP'; 'TPC'; and 'SBC' and a dependent variable – in the current research: students following teacher-directions ('On-task%') was multiple regression (Tabachnick and Fidell, 2012).

If co-linearity - or indeed, any other significant association - had been established between the 2 exponentially distributed teacher-feedback variables: 'SBP' and 'TPC'; and 'On-task%', or significant violation of homoscedasticity assumption (ibid.) then this decision would have required a further review, but this was not evident on inspection of scatter plots of distributions.

4.14 Hierarchical multiple regression analysis (HMRA) of UK27 teacher-feedback variables and 'On-task%'

Hierarchical multiple regression analysis (HMRA) – a variant of multiple regression (Palint, 2010) - was used to assess the relative predictive relationship between 5 control measures ('TPP'; 'SBP'; 'TPC'; 'SBC'; 'INX%') and the dependent variable: students following teacher-directions, ('On-task%'). The initial procedure using SPSS indicated that a particular case in the UK27 dataset (when N= 107) was causing violation of the procedure

by exceeding the critical value of the Mahalanobis distance for 5 variables (k= 5) of 20.26 for n>100, $p < .05$ (Tabachnick and Fidell, 2012). By removing the *one* identified outlier-case from UK27, as Tabachnick and Fidell advise for HMRA, the maximum Mahalanobis distance decreased (k=5) from 21.72 to 14.02. The sample using 5 independent variables as control measures for HMRA was still sufficiently large (N=106) with up to a 5% reduction of the dataset being judged as acceptable (ibid.).

Whilst the means of the 5 variables being considered as control measures with HMRA were more valid than medians in each case, (Pett, 2016), the exponential distributions, small size of means, large number of zero scores and exponential distributions (Tabachnick and Fidell, 2012) suggested caution about the inclusion of: 'SBP' and 'TPC' in the HMRA but results subsequently revealed that their influence in any of the 6 models was negligible, probably because of the high number of zero scores.

4.15 The hierarchical order of variable inclusion for HMRA of the predictive relationship of 5 UK27 independent teacher verbal-feedback variables when considering the distribution of the dependent variable 'On-task%' was: Model 1 – 'SBC'; Model 2 - add 'TPC'; Model 3 – add 'SBP' and 'TPP'; Model 4 – add 'INX%'; Model 5 – subtract 'SBC' and 'TPC'; Model 6 – subtract 'INX%':

Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**
Single most predictive independent variable	*Critical teacher-feedback only*	*Praising and critical teacher-feedback*	*Comprehensive teacher verbal feedback model*	*All positive + teacher-feedback + neutral talk*	*Only praising feedback*
SBC	SBC	SBC	SBC	SBP	SBP
	TPC	TPC	TPC	TPP	TPP
		SBP	SBP	INX%	
		TPP	TPP		
			INX%		

Table 12 – Hierarchical multiple regression: order of models and UK27 teacher verbal feedback variables

4.16 'SBC' accounted for 13.8 % of the variance (F (1, 104) = 16.63, $p < .001$) of 'On-task%'. 'TPC', when added in accounted for an additional 2% (F (2, 103) = 9.68, $p < .001$)

meaning that *critical* teacher-feedback as measured by these variables accounted for 15.8% of the variance.

When added in, 'SBP' and 'TPP' accounted for an additional 3.7% (F (4, 101) = 6.12, p < .001) of the variance; and 'INX%' for only an additional .1% (F (5, 100) = 4.89, p < .001); so the total variance of 'On-task%' accounted for by the comprehensive teacher verbal feedback model - Model 4, that included the 4 teacher-feedback variables, 'TPP' ($\mathcal{B}$[25]= .176 p= .075); 'SBP' ($\mathcal{B}$= -.129 p= .177); 'TPC' ($\mathcal{B}$= .135 p= .166); 'SBC' ($\mathcal{B}$= -.451 p< .001) plus the measure of total observed teacher-talk, 'INX%' ($\mathcal{B}$= .036 p= .696) was: 19.6%. Only the contribution of 'SBC' was individually significant (p< .05).

'SBC' and 'TPC' were subtracted to create Model 5 which consisted only of positive and / or neutral teacher-feedback: 'TPP'; 'SBP' and 'INX%'. Model 5 accounted for only 2.3% of the variance (F (3, 102) = .816, p< .5); and when 'INX%' was subtracted to create Model 6, the effect was negligible. Model 6 - 'TPP' and 'SBP' accounted for 2.3% of the variance (F (2, 103) = 1.233, p< .5) also.

The $\mathcal{B}$ (beta) statistic provided a relative measure for each predictor / control measure of their relative contribution to the comprehensive model; in order of importance within the UK27 dataset – the largest $\mathcal{B}$ is listed first: 'SBC'; 'TPP'; 'TPC'; 'SBP'; and 'INX%'. It is evident that 'SBC' provided by far the biggest single contribution in respect of accounting for the total variance in 'On-task%' within the UK27 dataset. It is also important to note that the *direction* of contribution as indicated by a negatively-signed $\mathcal{B}$ in the cases of both 'SBC' and 'SBP' is *negative* - meaning that *less* of both is predictively associated with *more* 'On-task%'.

4.17 UK27 teacher-feedback and 'On-task%' correlations

Correlational statistics are used here to provide supportive secondary information for associative relationships between variables and not primary arguments about distribution predictability. Correlation uses a transformational technique of standardising variable measurements in order to enable the possibility of a comparison of like and with like (Baguley, 2010). This ensures that any conclusions drawn are weaker than when using multiple regression and the $\mathcal{B}$ coefficient.

[25] $\mathcal{B}$ (Beta) = Standardised multiple regression coefficient produced by SPSS.

The next section investigates further the UK27 distributions of the four variables; 'TPP'; 'TPC'; 'SBP'; and 'INX%'; and their non-predictive relationship (as indicated by HMRA) with the distribution of the dependent variable 'On-task%' visually using scatter-plots and correlational statistics.

4.18 A *positive* correlation *expected* from the literature review and thus expressed *a priori* for the current research in Hypothesis #1 (see Chapter 1) was that 'On-task%' and 'TPP' would be significantly associated – meaning that teachers who used more positive feedback comments would be teaching classes where students would be more likely to be following their directions. There was no significant evidence that this was the case in the UK27 secondary school dataset from the parametric correlational analysis.

Investigating further the distributions of the variable 'TPP' and its less predictive relationship with the distribution of the dependent variable 'On-task%' visually using scatter-plots, the following 'TPP' graph shows clustering of scores in the left-upper quadrant:

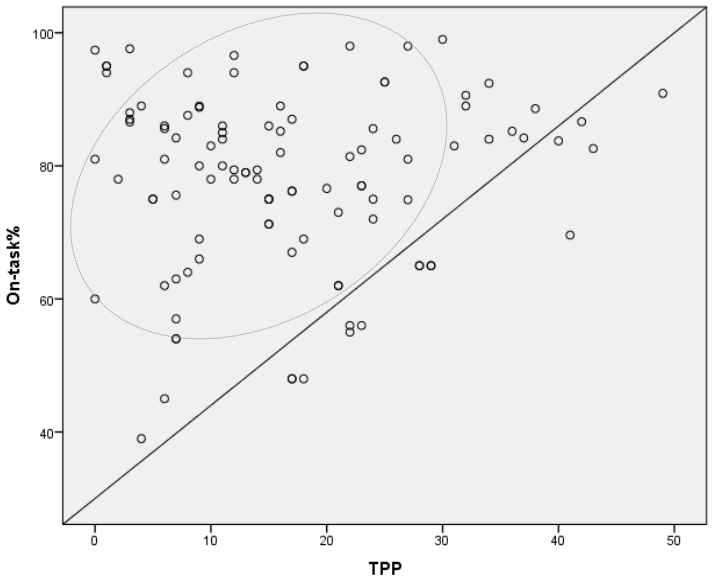


Figure 16 – Scatter plot showing UK27 'TPP' (Task Performance Positive) and its association with students following teacher-directions ('On-task%') (N= 106)

This distribution pattern is similar to that reported in the author's primary study (Apter, *et al.*, 2010, p. 156). On the face-validity of the scatter plot data, the distribution suggests that the association between the 2 variables is probably being affected by other unplotted variables. The clustering represented in the top-left quadrant suggests that tallies of between

0 and 30 ‘TPP’ comments were more associated with higher ‘On-task%’ scores of between 50% and 100% (marked with oval boundary); and that no very-low ‘On-task%’ figures (less than 50%) were associated with low / moderate rates of ‘TPP’ (tallies of less than 17 ‘TPP’ feedback comments. This observation might require further investigation.

As reported above, the distribution of ‘On-task%’ was approximately normal but ‘TPP’ less so. Spearman’s non-parametric rank correlation coefficient was insignificant (*rho*= .013, N= 106, *p*= .895, 2-tailed). This was unexpected and contrasts with the results from the author’s primary school research (Apter, *et al.* 2010, p.155) where a *positive* correlation was calculated (*r*= .345, *p* <.01, two-tailed).

4.19 As noted above, the distribution of ‘TPC’ data points approximately followed an exponential best-fit line with many tallies between 0 and 1 (see 4.10 above).

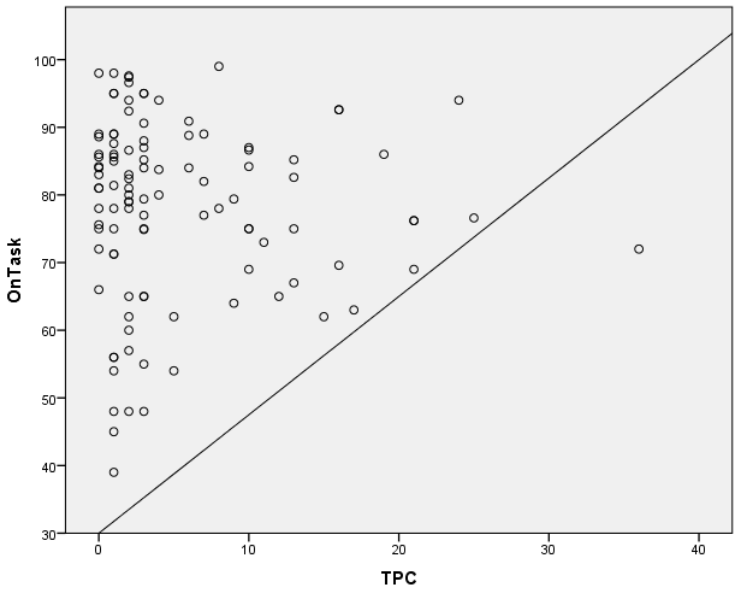


Figure 17 – Scatterplot showing UK27 ‘TPC’ (Task Performance Criticism) teacher comments and the association with students following teacher-directions (‘On-task%’) (N=106)

The non-parametric correlation between the two variables using Spearman’s product-moment correlation coefficient was insignificant (*rho*= .058, N= 106, *p*= .554, 2-tailed).

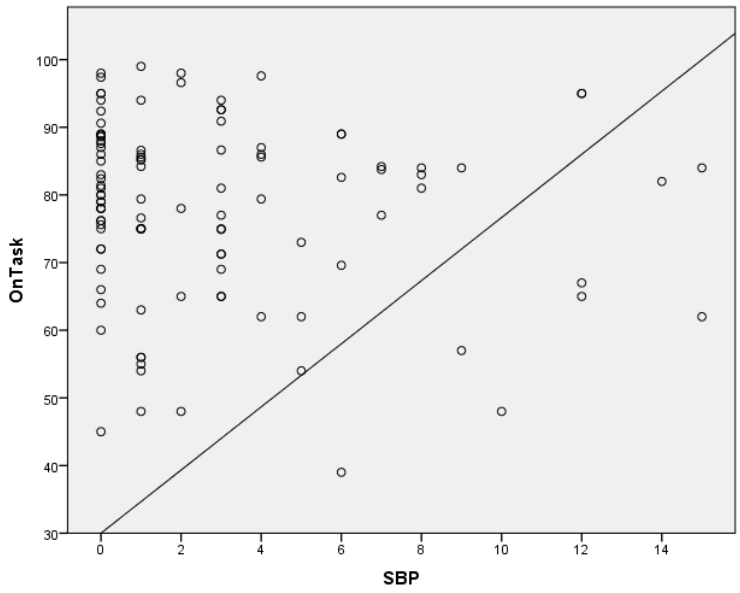


Figure 18 – Scatterplot showing UK27 'SBP' (Social Behaviour Positive) teacher comments and the association with students following teacher-directions ('On-task%') (N= 106)

As with TPC, the distribution of 'SBP' data points also followed approximately an exponential best-fit line with many tallies between 0 and 1 on the X axis. The non-parametric correlation between the two variables using Spearman's product-moment correlation coefficient was insignificant (*rho*= .125, N= 106, *p*= .201, two-tailed).

4.20 A *positive* correlation *expected* from the literature review and thus expressed *a priori* for the current research in Hypothesis #2 (see Chapter 1) was that 'On-task%' and 'INX%' would be significantly associated – meaning that teachers who spoke more as an integral element of their teaching would be teaching classes where students would be more likely to be following their directions. There was no significant evidence that this was the case in the UK27 secondary school dataset, either from the HMRA (4.16 above). This was unexpected and contrasts with the results from the author's primary school research (Apter, *et al.* 2010, p.155) where a significant *positive* correlation was calculated (*r*= .168, *p* =.05, two-tailed).

The following 'INX%' graph shows some clustering around the centre of the chart:

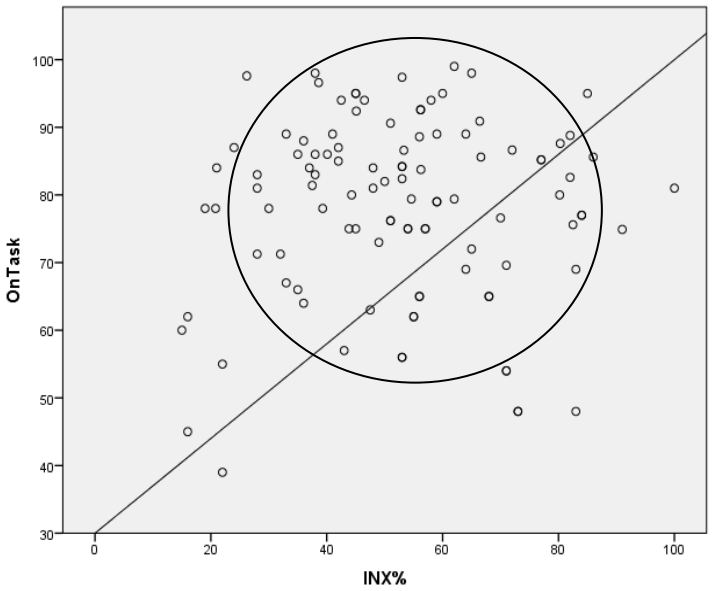


Figure 19 – Scatterplot showing UK27 'INX%' neutral teacher-talk and its association with students following teacher-directions ('On-task%') (N=106)

As noted above, the distributions of 'On-task%' and 'INX%' were near normal but the parametric correlation between the two variables using Pearson product-moment correlation coefficient was also insignificant (r= .040, N= 106, p= .681, 2-tailed).

4.21 In interpreting these results - either the predictive significance of the $\mathcal{B}$ coefficient when using HMRA, or the parametric *r*, or the non-parametric *rho* correlation - an important consideration was that the *MICRO* observation tool cannot be used to identify or analyse a direct correspondence between any particular student following teacher-directions and a teacher's feedback comments that might have specifically addressed their behaviour.

Thus the strongly predictive relationship (according to HMRA) between, for example, 'SBC' and the dependent variable 'On-task%' is not a causal one in either direction. For example: social behaviour criticism comments ('SBC') cannot be assumed to have *caused* episodes of specifically observed students following a teacher's directions ('On-task%') because the observations of each phenomenon were completed each time in sequence rather than simultaneously (see Chapter 3). The 'SBC' variable can only be used to predict the *distribution* of the 'On-task%' scores obtained during an observation (see Chapter 6 for a fuller discussion of these matters).

4.22 UK27 teacher-feedback type ratios

In common with findings from earlier research, (Wheldall *et al.,* 1989; and see Chapter 2), UK27 teachers comments critical of students' behaviour 'SBC' were used more than 'SBP' and 'TPC' but not as frequently as 'TPP'. The next section examines the significance of these differences within the UK27 dataset and whether Hypothesis #3 (Chapter 1) was supported.

The means of the 2 variables: 'TPC' (Task Performance Criticism) and 'SBP' (Social Behaviour Positive) in the UK27 dataset are much lower than the means of the other 2 variables representing teacher verbal feedback 'TPP' (Task Performance Positive) and 'SBC' (Social Behaviour Criticism). The following ratio table provides means and a comparison of ratios of the 4 types of teacher-feedback comments ('TPP'; 'SBP; 'TPC' and 'SBC'):

<table>
<tr><td>TPP Mean= 16.78
SD= 11.083
N = 106</td><td colspan="2">TPP / SBP 5.81 : 1</td><td>SBP Mean= 2.89
SD= 3.72
N = 106</td></tr>
<tr><td>TPP / TPC
3.28 : 1</td><td>TPP / SBC
1.37 : 1</td><td>TPC / SBP
1.77 : 1</td><td>SBP / SBC
1 : 4.22</td></tr>
<tr><td>TPC Mean= 5.11
SD= 6.02
N = 106</td><td colspan="2">TPC / SBC 1 : 2.39</td><td>SBC Mean= 12.21
SD= 9.87
N = 106</td></tr>
</table>

Table 13 - Ratio table of UK27 teacher-feedback means: 'TPP'; 'SBP; 'TPC'; and 'SBC'

A one-way repeated measures ANOVA was conducted to compare the means of teacher-feedback comments in 4 categories: 'TPP'; 'SBP; 'TPC'; and 'SBC'. There was a significant effect for teacher-feedback comments, Wilk's lambda, $\Lambda = .31$, (F (3, 103) = 75.10, $p<.005$), multivariate partial *eta* squared = .69. Pairwise comparisons revealed that each of the 4 category tallies are significantly different from each of the 3 other category tallies, $p< .05$.

Hypotheis #3 is therefore supported: teachers in secondary schools use significantly more *positive* verbal feedback, e.g. praising comments, to *encourage* academic-task-focussed

behaviour and product; and significantly more *negative* verbal feedback, e.g. critical comments, to discourage undesired social-behaviour ('Off-task' behaviour); and use significantly less *positive* verbal feedback to encourage students to follow their social-behavioural (non-academic) directions; and significantly less *negative* verbal feedback to discourage undesired, incorrect or inaccurate academic-task-focussed behaviour or product.

4.23 UK27 other teacher-feedback correlations

An *expected* correlation (following the results from the HMRA reported above) more fully revealed by a statistical exercise using Pearson's product-moment correlation coefficient was the finding of a negative medium correlation between 'SBC' and 'On-task%' (r= -.371, N= 106, significant at the .01 level, 2-tailed).

This confirmed the implication of the HMRA that the direction of the predictive relationship between the two variables meant that the less a teacher was observed to use social behaviour criticism ('SBC'), the more students would be observed to be following teacher-directions ('On-task%') in the classroom.

Whilst this finding was *post hoc* and not presented previously as a research hypothesis (in Chapter 1), there would appear to be an obvious face-validity to this finding - suggesting that a class that is generally well-behaved and following teacher-directions ('On-task%') are less likely to *provoke* social behaviour criticism ('SBC') from the teacher (see Chapter 6).

4.24 There were a number of additional correlations that were of *post hoc* interest evident from this phase of the statistical analysis of the UK27 dataset. 'SBC' was significantly *positively* correlated with 'TPP' (r= .195, *rho*= .208, p< .05, 2-tailed); and with 'TPC' (r= .357, *rho*= .351, p< .01, 2-tailed) meaning that teachers who used critical social behaviour feedback were also more likely to use 2 other forms of verbal feedback: praising and critical comments directed towards students' academic behaviour. A similar relationship was identified between 'SBP' and 'TPP' (r= .282, *rho*= .370, p< .01, 2-tailed).

In respect of the total amount of talking a teacher uses in the course of teaching an academic lesson ('INX%'), there was only one correlation with any teacher verbal feedback type ('TPP'; 'TPC'; 'SBP'; 'SBC'): a weak *positive* correlation using Pearson's product moment coefficient between 'INX%' and 'TPP' (r= .193, p< .05, 2-tailed) in the UK27 dataset. This result is reported here for the sake of completeness and as a contrasting comparator with the other more significant correlations reported here. The correlation was not replicated, *either*

non-parametrically using Spearman's *rho;* nor in the other dataset, UK1 (Chapter 5; Discussion).

4.25 UK27 contextual variables descriptives

HMRA was also used to assess the relative predictive relationship between 9 of the 10 contextual factors: 'Free school meals %'; 'School roll'; 'Local type'; 'AM / PM'; 'Class size'; 'Adults present'; 'Years' experience'; 'Teacher gender'; 'Year group', and the dependent variable: students following teacher-directions, ('On-task%'). The 10th independent variable: 'Subject' was excluded because it was a non-dichotomous categorical variable with 7 levels (academic subjects): English, maths, science, history, geography, other (including media, and personal development), and modern languages. It was not possible to usefully recode these as pairs of dummy variables[26]. In contrast, whilst technically a categorical variable, 'Year group' was treated as a continuous variable for the purpose of HMRA. This was justified because of the close relationship between year-group and age, meaning that 'Year group' could be treated as a continuous variable.

4.26 The maximum Mahalanobis distance for 9 variables (k= 9) was 27.88 for n>100, p < .01 (Tabachnick and Fidell, 2012). One case (*MD*=50.22) exceeded this critical value and was removed from the dataset. By removing an *additional* identified outlier-case from UK27, as Tabachnick and Fidell advise for HMRA, the maximum Mahalanobis distance decreased for 9 variables from 50.22 to an acceptable 27.67. The sample was still sufficiently large (N=105) with a less than 1% reduction of the dataset being judged as acceptable (ibid.). The means of the 9 variables being considered as control measures with HMRA were more valid than medians in each case, (Pett, 2016). 'Locality type' (scores 1 or 2, mean= 1.43, SD= .497); 'Teacher gender' (scores 1 or 2, mean= 1.52, SD= .50); and 'AM or PM' (scores 1 or 2, mean= 1.31, SD= .465) were dichotomous and therefore admissible for HMRA despite violating normality (ibid.).

Distributions of the other 6 variables failed the Kolmogorov-Smirnov test. This is unsurprising given the nature of these contextual variables and the large (n>100) sample size.

[26] When 'Subject' was recoded as dummy variables, this created 7 additional variables and many more outliers, many of which exceeded the critical value of the Mahalanobis distance. Removal of 2 additional cases and a significant increase in the number of predictors further compromised the veracity of the HMRA that was attempted and it was abandoned.

For example: ‘Roll’ and ‘Free school meals%’ are anomalously distributed but with similar distribution patterns that are negatively correlated (r= -.240, p< .05, 2-tailed) with an evident demographic relationship.

4.27 ‘Year group’ had left-skewed curve which could be described as a self-selection effect (Ziliak and McCloskey, 2008) with a mean of 8.66 and a 5% trimmed mean of 8.58. Skewness was measured at S= .636 (SE= .235); and kurtosis at K= -.738 (SE= .465):

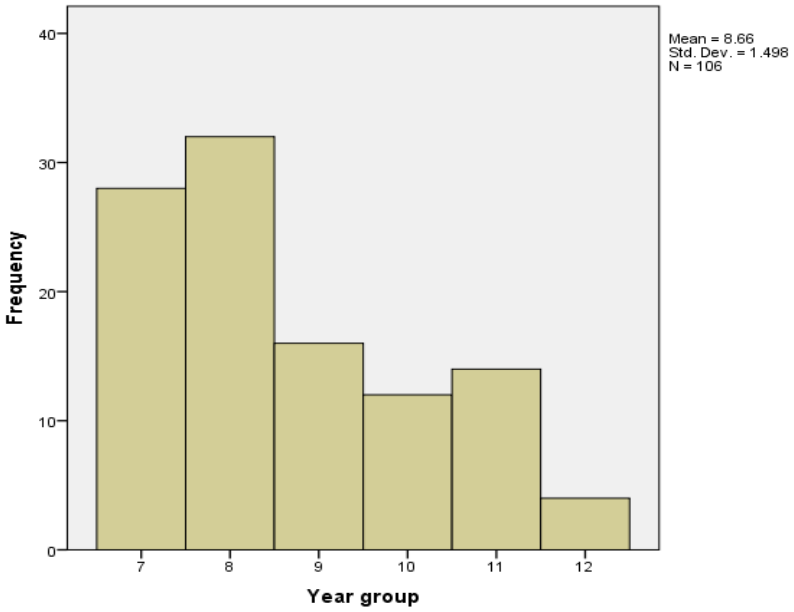


Figure 20 – Histogram showing distribution of UK27 ‘Year group’ frequencies (N=106)

As can be seen from this histogram, Year 7 and 8 classes were selected most frequently by observers for the current research probably because they were probably the most convenient classes to select. This tails off to much lower frequencies for the GCSE examination years 10 and 11. There were just 2 observations offered of Year 12 classes where students had begun study for their A Level examinations and none for Year 13, the A Level examination year.

4.28 The distribution of the UK27 ‘Years’ experience’ frequencies accords to a left-skewed high peaked normal distribution with a mean of 8.20 and a 5% trimmed mean of 7.77. Skewness was measured at S= 1.232 (SE= .235); and kurtosis at K= 1.149 (SE= .465):

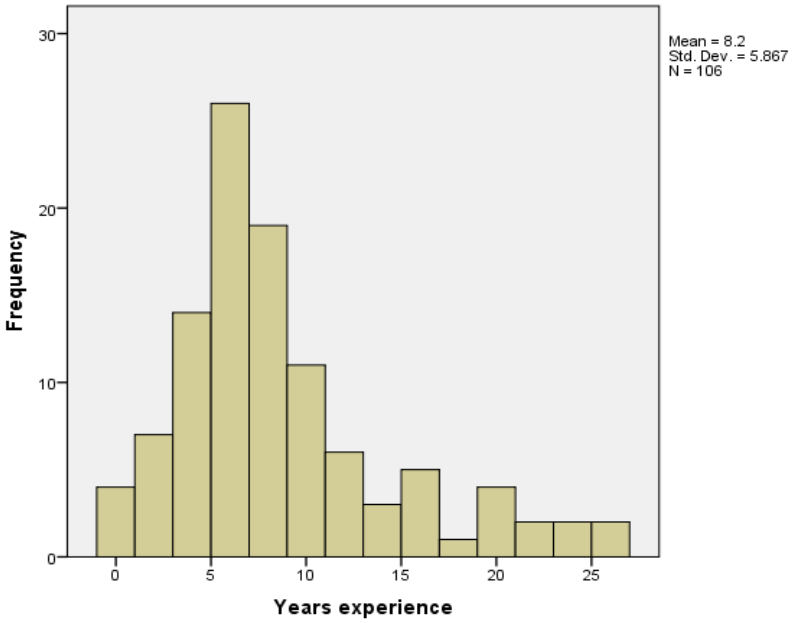


Figure 21 – Histogram showing distribution of UK27 'Years' experience' frequencies (N=106)

This distribution presents as an accurate analogue of the teacher-retention figures for all UK schools (DfE, 2015b).

4.29 'Adults present' in the class was correctly anticipated to be an exponentially shaped distribution for obvious reasons:

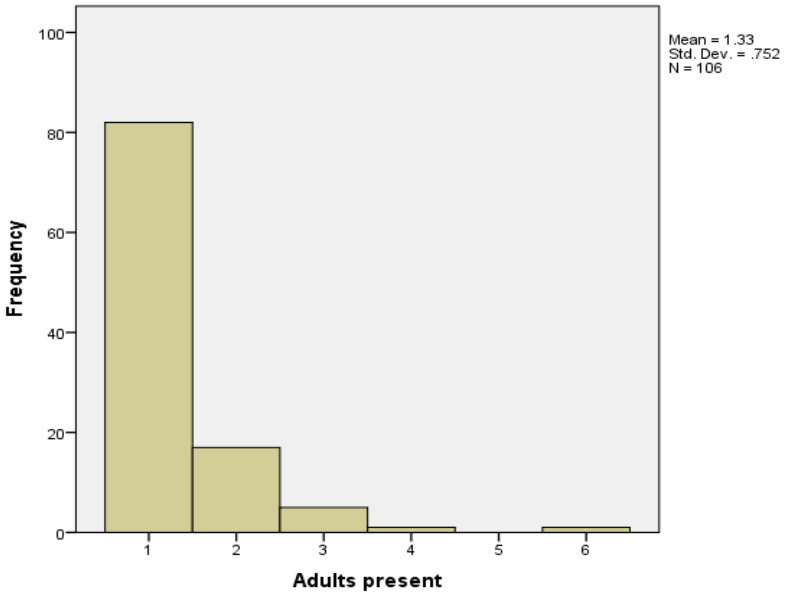


Figure 22 – Histogram of UK27 'Adults present' in class (N= 106)

The anomalous single case where there were 5 adults in addition to the class-teacher supporting students with special needs was a Year 9 science lesson where students were completing an experiment with hazardous apparatus in small groups.

4.30 A speculation with the distribution pattern of the UK27 'Class size' contextual variable represented in the histogram below is that there are 2 separate distribution patterns overlaid upon each other – predominantly large 20 to 30 student groups for core subjects, English and maths in years 7 through to 9 dominate, and then smaller specialised groups dominate for more specialised lessons in years 10 and 11. The overlay of the two or more patterns could be judged to have created the cumulative chaotic pattern presented here:

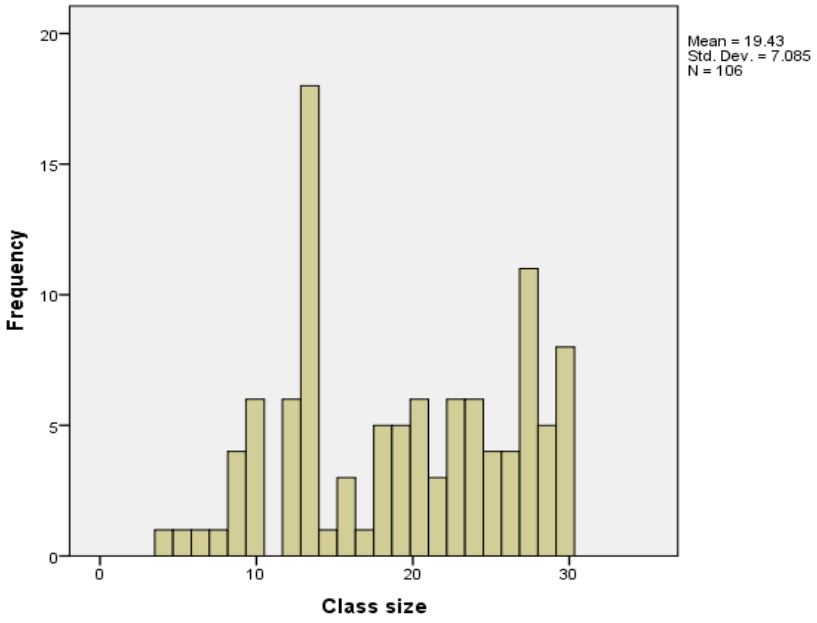


Figure 23 – Histogram of UK27 'Class size' (N= 106)

A case-numbered scatter-plot suggests that this might be the case.

There are a number of clusters evident but the various effects that might be represented by the clusters are complex and would require considerable further investigation for any hypothesis to be generated and supported and there is no current indication that such an emergent hypotheses would be appropriate within the province of the current research:

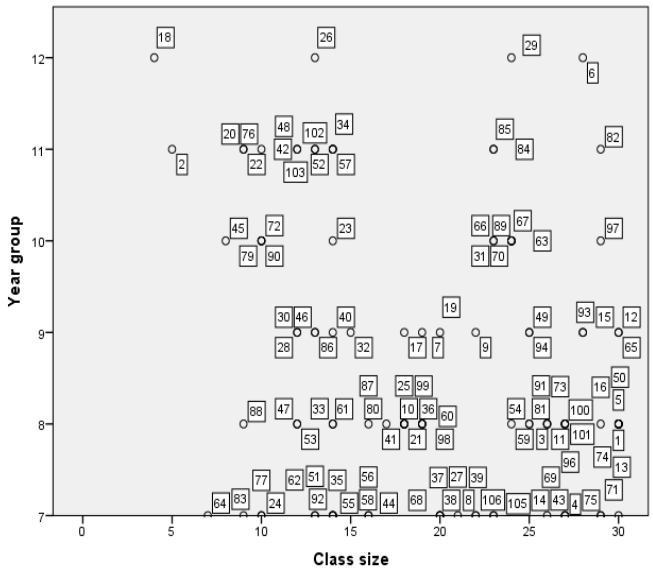


Figure 24 – Case-numbered scatter plot of UK27 'Class size' by 'Year group' data-points (N= 106)

4.31 It is not appropriate to treat the array of curricular 'Subjects' and their frequencies as a normally-distributed variable and it will not be included in the UK27 HMRA of 9 contextual variables (below). 'Subjects' is a category variable, the levels (or categories) of which are separate curricular subjects. Whilst each curricular area included in 'Subjects' is described as being delivered as a classroom-based academic lesson, 'Other' is a catch-all category that includes more informal lessons such as pastoral guidance and citizenship that do not culminate in an examination. Some lessons were not included in the study, for example: P.E.; I.T.; and art.

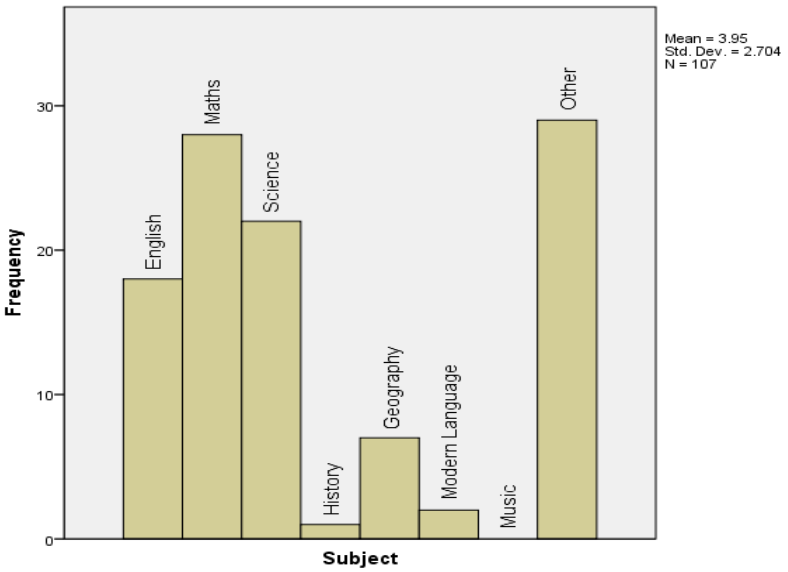


Figure 25 – Histogram showing distribution of curricular UK27 'Subjects' frequencies (N= 106)

The histogram provides a representation of the distribution of curricular subject-lessons that were observed. The distribution violates an assumption of normality.

4.32 UK27 HMRA of contextual variables

The hierarchical order of variable inclusion for HMRA of the predictive relationship of 9 UK27 independent contextual variables when considering the distribution of the dependent variable 'On-task%' was: Model 1 – 'Teacher gender' and 'Years' experience'; Model 2 - add 'Year group'; Model 3 – add 'Class size' and 'Adults present'; Model 4 – add 'Locality type', 'Free school meals' and 'Roll'; Model 5 – add 'AM or PM':

Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**
Two most predictive teacher-specific contextual variables	*Hypothetically next most predictive contextual variable*	*Add human context: class size plus no. of adults*	*Comprehensive teacher verbal feedback model*	*All positive + teacher-feedback + neutral talk* *Only praising feedback*
Years' experience Teacher gender	Years' experience Teacher gender	Years' experience Teacher gender	Years' experience Teacher gender	Years' experience Teacher gender
	Year group	Year group	Year group	Year group
		Class size Adults present	Class size Adults present	Class size Adults present
			Locality type Free school meals Roll	Locality type Free school meals Roll
				AM or PM

Table 14 – Hierarchical multiple regression: order of introduction of models and UK27 contextual variables

4.33 Consisting of two teacher-specific contextual factors: 'Years' experience' and 'Teacher gender', Model 1 accounted for 17.2 % of the variance (F (2, 102) = 10.56, p < .001) in 'On-task %'.

Model 2 - when 'Year group' was added in to the HMRA, accounted for an additional 2.6% of the variance (F (3, 101) = 8.31, p < .001). Model 3, when 'Class size' and 'Adults present' were added in, accounted for only an additional .8% of the variance (F (5, 99) =

5.12, $p < .001$). Model 4 - when 'Locality type'; 'Free school meals %'; and 'Roll' were added in to the HMRA, accounted for an additional 6.9% of the variance (F (8, 96) = 4.55, $p < .001$). Model 5 - when 'AM or PM' was added in to the HMRA, accounted for an additional .1% of the variance (F (9, 95) = 4.03, $p < .001$).

Model 5 - the comprehensive contextual model which included 9 contextual variables (excepting the inclusion of 'Subject'), accounted for a total 27.6% of the variance in the 'On task %' distribution. Model 5 consisted of (in order of predictive importance – the most influential variable first): 'Years' experience'($\mathcal{B}$= .382 $p < .001$); 'Free school meals %' ($\mathcal{B}$= .313 $p < .005$); 'Teacher gender' ($\mathcal{B}$= -.189 p= .051)[27]; 'Year group' ($\mathcal{B}$= -.181 p= .065); 'Locality type' ($\mathcal{B}$= .177 p= .135); Adults present' ($\mathcal{B}$= .075 p= .471); 'Roll' ($\mathcal{B}$= -.054, p= .646); 'Class size' ($\mathcal{B}$= -.031 p= .77); ' 'AM or PM' ($\mathcal{B}$= .39, p= .673). Only the contributions of 'Years' experience' and 'Free school meals' were individually significant ($p < .05$).

4.34 Other correlations in the UK27 dataset involving contextual variables

There were a number of emergent un-hypothesised contextual-variable correlations in the UK27 dataset revealed by using Spearman's rank order coefficient, *rho*. The dataset is unequally representative in terms of year groups. As correlations were not hypothesised and the significant correlations do not have relevance for the research questions of the current research (Chapter 2), no further explanations has been attempted. Correlations are included here for the sake of completeness but should be considered with caution.

In the UK27 dataset, 'Locality type' was correlated with 'INX%' (*rho*= .233, N= 106, $p < .05$), 'TPC' (*rho*= .308, N= 106, $p < .01$), 'Free school meals%' (*rho*= -.675, N= 106, $p < .01$) and 'Class size' (*rho*= .369, N= 106, $p < .01$); 'Free school meal%' with 'SBP' (*rho*= -.275, N= 106, $p < .01$), 'TPC' (*rho*= -.313, N= 106, $p < .01$), 'Adults present' (*rho*= .244, N= 106, $p < .05$) and 'Class size' (*rho*= -.459, N= 106, $p < .01$); 'Year group' with 'TPP' (*rho*= -.325, N= 106, $p < .01$), 'Adults present' (*rho*= -.232, N= 106, $p < .05$), and 'AM or PM' (*rho*= .216, N= 106, $p < .05$); 'Adults present' with 'TPP' (*rho*= -.208, N= 106, $p < .05$); 'Class size' with 'SBP'(*rho*= .263, N= 106, $p < .01$); and 'AM or PM' with 'TPP' (*rho*= -.203, N= 106, $p < .05$).

[27] Negatively signed statistic relates to dichotomous variable, male= 1, female= 2; therefore, ($\mathcal{B}$= -.189 $p < .06$) indicates that male teachers are more predictively associated with higher 'On-task %' but at an insignificant level ($p > .05$).

4.35 UK27 teacher-feedback by year group

Looking provisionally at the relationship between 'Year group', 'On-task%' and the 4 teacher verbal feedback types: 'TPP'; 'TPC'; 'SBP'; and 'SBC'; and the neutral teacher verbal behaviour percentage, 'INX%', it is instructive to use a line-graph to visually depict trends in the UK27 dataset, year-group by year-group:

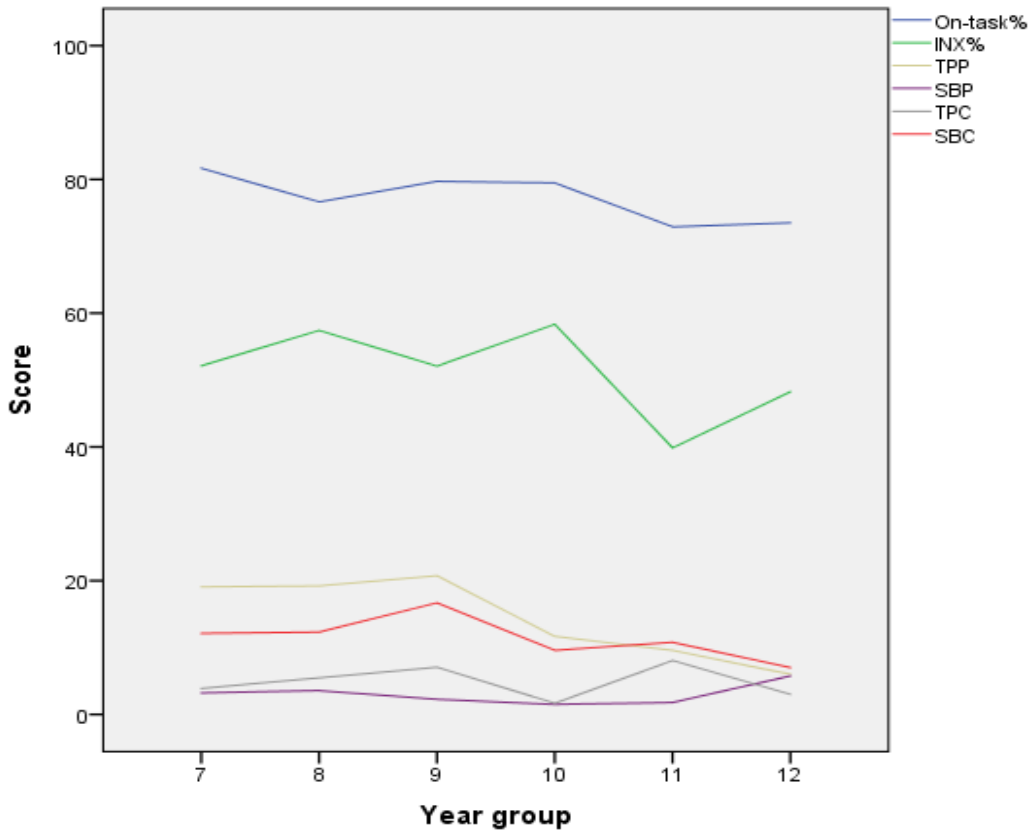


Figure 26 – line graph depicting the UK27 frequency of 6 variables – 'On-task%' and 5 teacher verbal behaviour variables by 'Year group'

It is important to be reminded that the top 2 lines: 'On-task%' and 'INX%', are both recorded on the graph as percentages in contrast to the 4 types of teacher-feedback variables, which are recorded on the graph as means of tallies of feedback-comments. The graph depicts how teachers use much more 'TPP' and 'SBC' than 'SBP' and 'TPC' in each year group but that the tally-means converge downwards to low means in Year 11 and 12. This reflects maybe the teacher-expectation that students are able to manage with less teacher support including verbal feedback as they mature as learners and progress year by year through their secondary school.

The graph suggests that the UK27 'On-task%' figure starts in Year 7 at 82% - nearly at the same level that was recorded in the author's primary school study (85%; Apter, *et al.*, 2010) but then declines to 72.5% by Year 11. The graph also suggest that 'TPP' and 'SBC' are used more by teachers in Years 7 and 8 but usage of both decreases year by year in a downward trend to Year 11 and 12. 'INX%' also appears to dip as Year 11 is reached.

Distilling this graph in to the relationships - year group by year group - of the means of 'On-task%'; 'TPP'; and 'SBC' which were suspected of being the most revealing of any covariance in the trends depicted by the graph, the following table was derived:

	Year group	Mean	SD	N
On-task%	7	81.67	9.634	28
	8	76.36	14.175	30
	9	81.83	13.5	15
	10	79.48	6.098	12
	11	72.92	18.051	14
	12	73.50	18.156	4
	Total	78.38	13.155	103
TPP	7	19.07	10.691	28
	8	20.07	12.357	30
	9	21.00	10.549	15
	10	11.67	7.738	12
	11	9.57	6.676	14
	12	6.00	7.3485	4
	Total	16.98	11.164	103
SBC	7	12.11	11.229	28
	8	11.90	8.942	30
	9	15.87	11.544	15
	10	9.58	7.489	12
	11	10.79	9.014	14
	12	7.00	7.616	4
	Total	11.92	9.828	103

Table 15 – Means and standard deviations by 'Year group' of 'On-task%'; 'TPP'; and 'SBC' in the UK27 dataset

What appear to be visible downward frequency-trends, particularly in respect of 'TPP' and 'On-task%' decreasing as students move up through the UK27 secondary school year-

groups is evident from an inspection of these means. However obvious this relationship might appear, it was not hypothetically envisaged when this research was designed (Chapter 1); therefore, two additional *post hoc* hypotheses are: that teachers in UK secondary schools use *less* positive feedback to older students in academic lessons; and older students appear significantly less likely to be following directions given them by teachers in academic lessons.

4.36 A one-way between-groups multivariate analysis of variance (MANOVA) was performed on the UK27 dataset to investigate 'Year group' differences upon classroom behaviour – both teacher and student. Rather than considering only the dependent variable used in the main part of this study, 'On-task%', 'TPP' and 'SBC' were added-in as combined dependent variables to provide a fuller picture of the interdependence and covariance of teachers' verbal feedback with students' 'On-task%' behaviour within the UK27dataset. 'TPP' and 'SBC' were chosen as being more dynamically indicative of teachers' verbal behaviour because they had been measured to be higher frequency and the most significantly associated variables with the variance of the distribution of 'On-task%'.

The independent variable was 'Year-group' which for the purpose of MANOVA was now treated as a categorical variable[28] with 5 groupings (Y7, Y8, Y9, Y10, Y11 and Y12). Preliminary assumption testing was conducted to check for normality, linearity, univariate and multivariate outliers, homogeneity of variance-covariance matrices, and multicollinearity, with one violation noted: 'On-task%' failed Levene's test of equality of error variances, ($F(5, 97) = 4.075$, $p < .05$) but the other two variables were non-violating ('TPP': $F(5, 97) = 1.658$, $p > .1$; 'SBC': $F(5, 97) = 1.064$, $p > .25$). By arithmetically transforming the variable 'On-task%'[29], all 3 variables passed the Levene test ('On-task%': $F(5, 97) = 1.75$, $p > .1$; 'TPP': $F(5, 97) = 1.658$, $p > .1$; and 'SBC': $F(5, 97) = 1.064$, $p > .1$).

There were statistically significant differences between 'Year group' on the combined dependent variables, $F(15, 263) = 1.97$, $p = 0.18$; Wilks' *lambda* = .745; partial *eta* squared= 0.93. When the results for the dependent variables were considered separately, the only difference to individually reach statistical significance with a Bonferroni-adjusted *alpha* level of 0.17, was 'TPP', $F(5, 97) = 4.15$, $p = .002$, partial *eta* squared= .176.

[28] For HMRA, 'Year group' had been treated as a continuous variable.
[29] 'On-task%'= (EXP('OnTask%'/10))/100

Both additional *post hoc* hypotheses are supported: that teachers in UK secondary schools do use *less* positive feedback to older students in academic lessons; and older students appear significantly *less* likely to be following directions given them by teachers in academic lessons.

4.37 UK27 teacher-feedback by subject

Whilst 'Subject' could not be included in the HMRA of UK27 contextual variables, the correlations between 'Subject' with the dependent variable 'On-task%' can be reported. Neither parametric (*r*) nor non-parametric (*rho*) correlational analysis of the relationship between 'Subject' dummy-variables, 'English', 'Maths' and 'Science' revealed any significant relationships. This means that within the UK27 dataset there was no evidence that teachers used significantly different patterns or amounts of verbal feedback and neutral verbal behaviour ('TPP'; 'TPC'; 'SBP'; 'SBC'; 'INX%') ; and that students did not follow directions to a more significant degree in any 1 of these 3 core subjects. Other subjects, such as 'Modern language', 'History' and 'Geography', were not used for this part of the investigation because the number of lessons observed was so small in each case.

Chapter 5 – UK1 Results - analysis of comprehensive set of classroom observations in one UK secondary school (UK1 dataset)

- **UK1 dataset descriptive statistics**
- **UK1 dependent variable ('On-task%')**
- **Graphs of UK1 teacher verbal feedback variables**
- **Hierarchical multiple regression analysis (HMRA) of UK1 teacher-feedback variables and 'On-task%'**
- **UK1 teacher-feedback and 'On-task%' correlations**
- **UK1 teacher-feedback type ratios**
- **UK1 teacher-feedback types by year-group**
- **UK1 teacher-feedback additional findings**
- **UK1 contextual variables descriptives**
- **UK1 HMRA of contextual variables**
- **Other correlations in the UK1 dataset involving contextual variables**
- **UK1 teacher-feedback by year group**
- **UK1 teacher-feedback by subject**
- **Comparison of UK1 dataset analysis of teacher-feedback means with UK27 teacher-feedback means**
- **Comparison of UK1 contextual means with UK27 contextual means**
- **Summative account of an Iterative Learning Conversation with UK1 school using a 3x pass Foucauldian thematic discourse analysis**

5.01 UK1 dataset descriptives

The UK1 dataset was compiled from *MICRO* data-summary sheets returned by 4 observers of 129 academic lessons[30] in one UK midlands city secondary school[31] catering for 980 students. The school was familiar and local to the professional province of the author. Each teacher in the school was observed twice, once in the morning and once in the afternoon. The school was identified as requiring 'special measures' by Ofsted at the start of the observation period (Chapter 6).

Whilst a full range of academic and non-academic lessons were observed, as with the UK27 dataset only observational data derived from academic ('*teach, talk and* write') lessons were used in this study.

A set of 7 observations were subsequently removed from UK1 for the HMRA and ANOVA statistical process reported in this chapter because the Mahalanobis distance calculated for

[30] English, maths, science, history, geography, and other - miscellaneous desk-based 'teach, talk and write' lessons, e.g. R.E., personal and social education – P.S.E., modern language, etc.
[31] The school was identified as requiring 'special measures' by Ofsted at the start of the observation period and this was in-part the reason for school's senior leadership team (SLT) inviting the author to complete this strand of the research in the school (see Chapter 5 – Discussion).

this data exceeded the critical value for 5 variables used for teacher verbal feedback analysis and their effects upon the dependent variable. These data were not subsequently re-included for other analyses included in the current research. A total of 122 observations were therefore used for the analysis of the UK1 dataset.

5.02 The UK1 secondary school in May, 2014 was classified as an inner-city secondary school ('Locality type'= 2), with 980 students on roll ('School roll'= 980), and 22.26% of students received free school meals ('Free school meals%= 22.26%). These 3 fixed UK1 data points are comparators with 3 contextual variables with similar names used in the UK27 dataset.

The following 13 variables (5 teacher verbal feedback variables, 7 contextual variables, and the singular dependent variable 'On-task%') were also included in the analyses of the UK1 dataset:

	Variable Name	***Description***	***Range***
d.v.*	On-task%	Following teacher-directions mean-time as %	0 - 100
Teacher verbal behaviour – independent variables	TPP	Task performance positive teacher-comment	0 – 25+
	SBP	Social behaviour positive teacher-comment	0 – 2+
	TPC	Task performance criticism teacher-comment	0 – 5+
	SBC	Social behavioural criticism teacher-comment	0 – 25+
	INX%	Neutral teacher-talk as a proportion of time %	0 – 100
	Year group	National curriculum year of class	7 - 12
	Teacher gender	Male / female	1= m. or 2= f.
	Years' experience	Number of years teacher has taught	Less than 1 to 25
	Subject	English, maths, science etc.	1,2,3,4,5,6*
	Adults present	Number of adults in classroom	1 – 4+
	Class size	Number of students in classroom	3 – 30+
	AM or PM	Lesson in morning or afternoon	1= a.m. 2= p.m.

*UK1 Additional Codes: d.v.= dependent variable, m.= male, f.=female, 1= English, 2= Maths, 3= Science, 4= History, 5= Geography, 6= Other

Table 16 - Key to UK1 data variable types

5.03 The following table displays the data limitations of the UK1 dataset in terms of simple descriptive statistics and minimum / maximum recorded values for 13 variables.

	Variable	*N*[32]	*Min.*	*Max.*	*Mean*	*SD*
d.v.	On-task%	122	34.80	100.00	77.06	14.44
Teacher verbal behaviour – independent variables	TPP	122	0	25	6.63	6.088
	SBP	122	0	2	.189	.43
	TPC	122	0	5	.861	1.208
	SBC	122	0	25	6.148	6.558
	INX%	122	8	100	31.814	15.884
	Year group	122	7	12	9.51	1.721
	Teacher gender	122	1	2	1.57	.497
	Years' experience	122	0	34	10.31	8.43
	Subject	122	1	6	3.87	1.945
	Adults present	122	1	4	1.52	.707
	Class size	122	3	30	17.19	7.646
	AM or PM	122	1	2	1.48	.502

Table 17 – UK 1 dataset, descriptive statistics

Classes and teachers were selected by observers as matching the research inclusion criteria provided in advance documentation (appended: Appendix; see also Chapter 3). The mean duration of UK1 lesson observations was 30.3 minutes (SD= .248).

5.04 The following figure is a graph depicting boxplots of 13 variables (N=122) used in the UK1 dataset:

[32] Originally n=129 after initial data-veracity sift; 7 observations were subsequently removed for the statistical procedures reported in this chapter because values exceeded Mahalanobis distance (see below.)

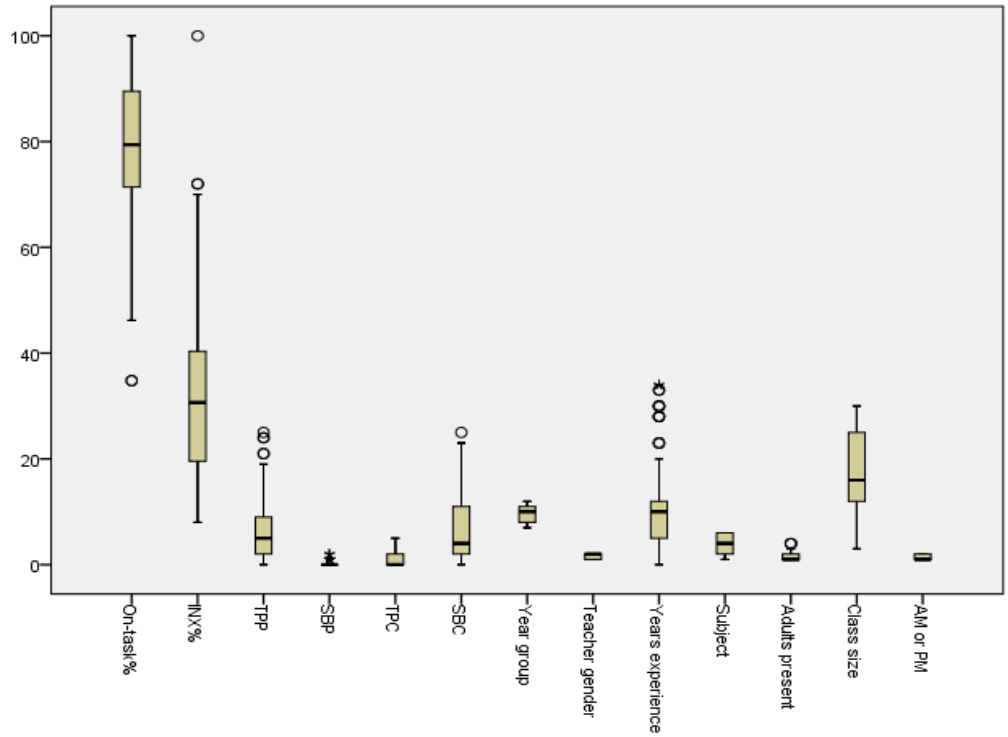


Figure 27 – Boxplots of UK1 variables included in the current research (N=122)

This graph serves to graphically represent data distribution centred on UK1 variable means, and any outliers. With 'On-task%' and teacher verbal-feedback: there were 2 lower-outliers from the boxplot of 'On-task%'; 3 upper-outliers from 'INX%'; 5 from 'TPP'; 3 from 'SBP' and 1 from 'SBC'. With school-context variables: there were 11 upper-outliers from the boxplot of 'Years' experience' and 3 from 'Adults present'.

In the case of 'TPC' and 'SBP' an exponential distribution and upward-outliers was expected from the review of previous literature: teachers have not previously been observed to make significant usage of these feedback-types (Chapter 2) and this was confirmed as a feature of the UK27 dataset (Chapter 4).

The smallest boxes represent the 2 dichotomous context variables: 'Teacher gender'; and 'AM or PM'.

5.05 UK1 dependent variable ('On-task%')

The distribution of the UK1 dependent variable: 'On-task%' is represented in the following bar graph:

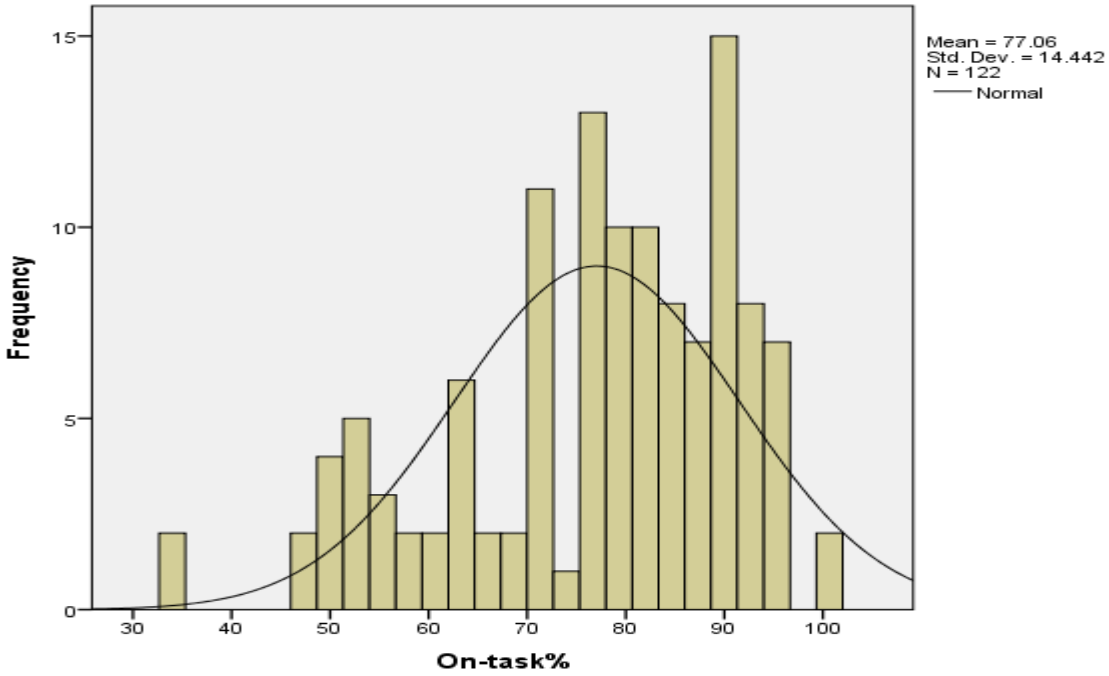


Figure 28 – Distribution of 'On-task%' (following teachers' directions) in UK1 dataset with 'best-fit' normal distribution curve (N= 122)

Testing the UK1 'On-task%' data for being normally distributed, the distribution failed the Kolmogorov-Smirnov test with D (122) = .12, p < .005. As with the UK27 dataset, this might appear to mean that parametric statistical analytics should not be used with the UK1 dataset but in establishing a predictive relationship between the distributions of one or more independent variables and the distribution of a dependent variable to support a hypothesis, parametric tests are more powerful than their non-parametric equivalents (Pett, 2016).

The 5% trimmed mean of the UK1 dependent variable ('On-task%) was satisfactory at 77.82% (ibid.). Skewness was measured at S= -.82 (SE= .219) which described the midpoint of the distribution being skewed to the right as is evident from the histogram above (Figure 28). Kurtosis, measured at K= .11 (SE= .435) suggested that the distribution was steeply peaked.

5.06 The Q-Q Plot (Figure 29) indicated normality whilst reflecting the previously noted caveat about skewness:

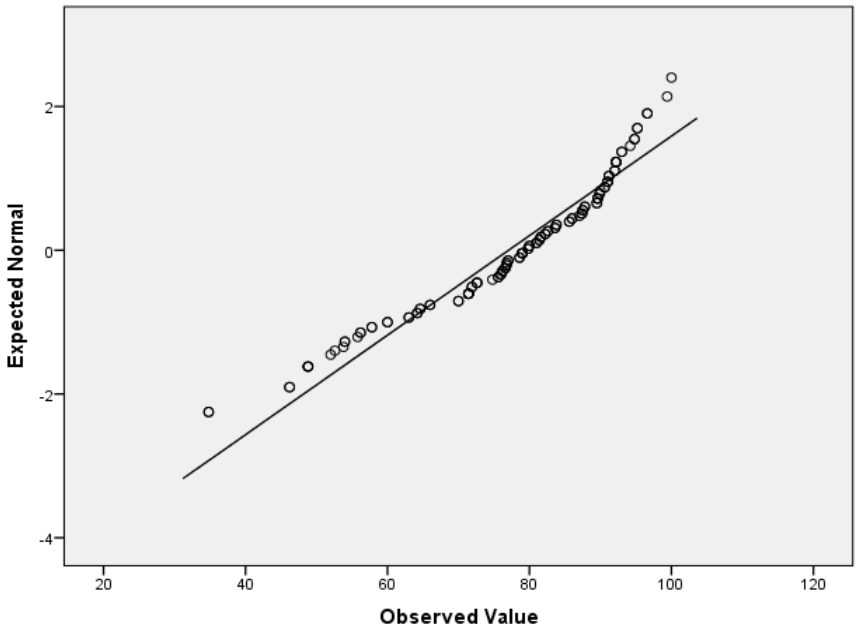


Figure 29 – Normal Q-Q Plot of UK1 'On-task%' variable

The UK1 sample was moderately large (N= 122); the mean was more valid than the median as a significant characteristic of the distribution; and the number of 'On-task%' outliers was small (N= 2), not too distant from the median and in the direction of the more complete leading tail - given the right-handed skew of the distribution[33].

5.07 The mean following teacher-directions ('On-task%') percentage at 77.06% for the UK1 secondary school including all year groups (7 to 12) was 1.09% lower than 78.15% calculated for the UK27 dataset, and 7.43% lower than the mean 'on-task' percentage of 84.49% obtained in the author's similarly large-scale 2010 primary school study (Apter, *et al.*, 2010; N= 141).

5.08 Graphs of UK1 teacher verbal feedback variables

The 8 graphs below are histograms and respective Q-Q Plots representing 4 types of teacher-feedback comments included in the UK1 dataset: 'TPP' (Task Performance Positive); 'SBC' (Social Behaviour Criticism); 'TPC' (Task Performance Criticism) 'SBP' (Social Behaviour Positive):

[33] As with the UK27 dataset, the UK1 dataset was also subjected to non-parametric tests particularly when a distribution violated parametric assumptions.

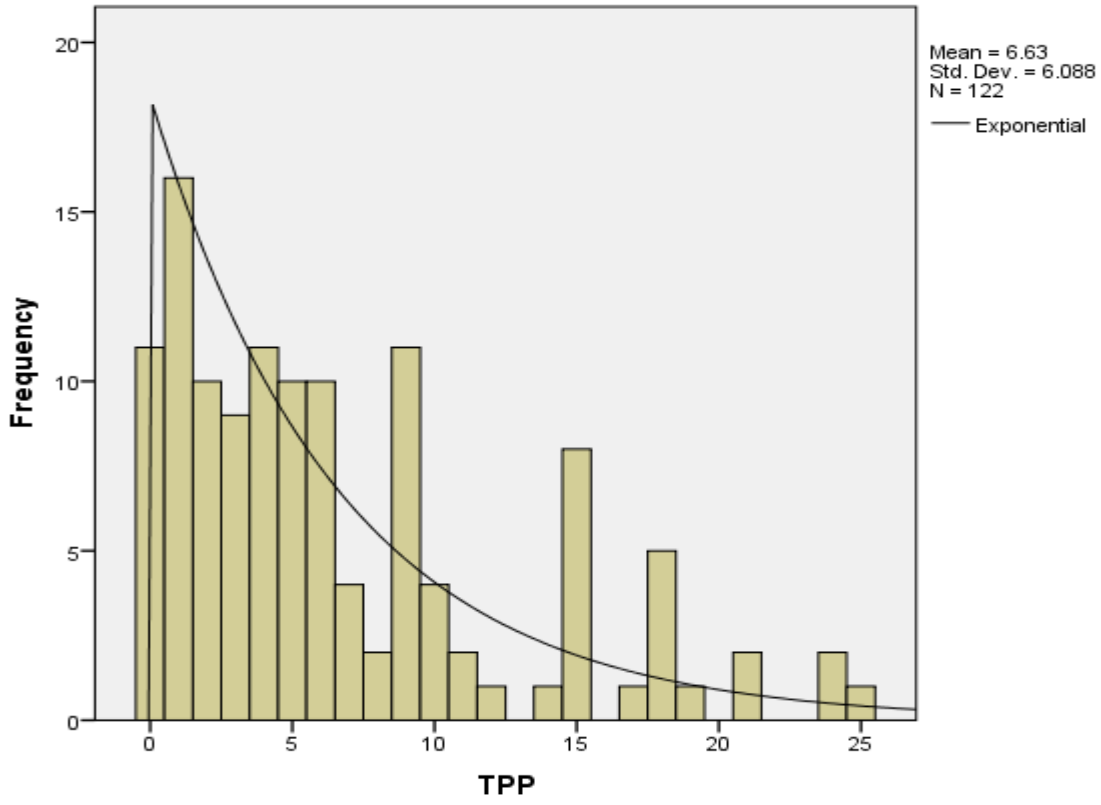


Figure 30 - Histogram representing UK1 'TPP' (Task-Performance Positive) teacher-feedback comments with best-fit exponential distribution curve (N=122)

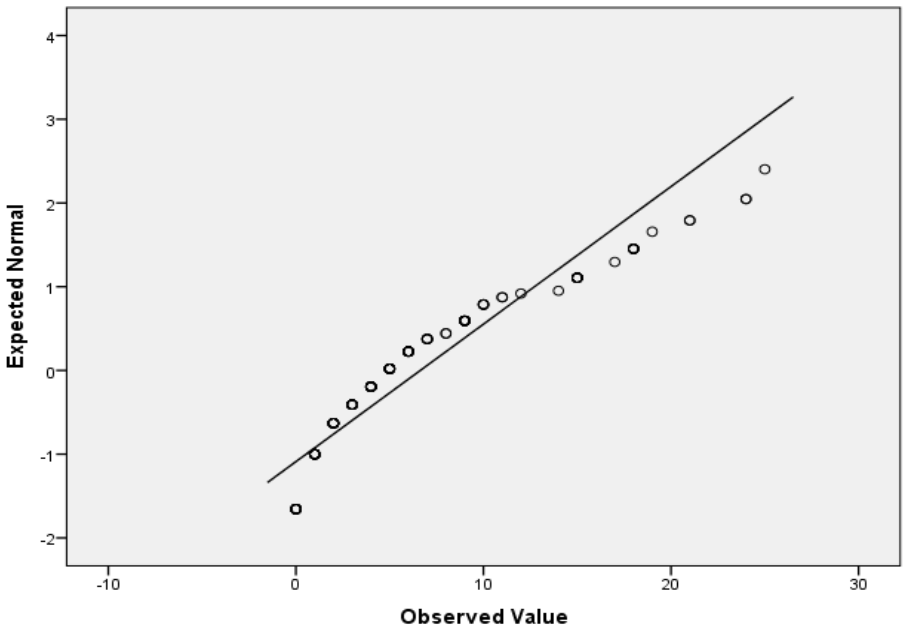


Figure 31 – Normal Q-Q Plot of UK1 'TPP'

The mean and distribution suggests that teachers used positive feedback-comments directed at students task performance ('TPP') at a much lower rate (mean= 6.68, N=122) – meaning that teachers in the UK1 school positively recognise students for academic work much less in comparison with the UK27 'TPP' figure (16.78, N=106). The histogram also demonstrates that the distribution of 'TPP' is more exponential than normal.

5.09 The following histogram illustrates the distribution of 'SBP' and the rates of positive feedback comments directed towards social behaviour that teachers used in academic lessons in the UK1 school:

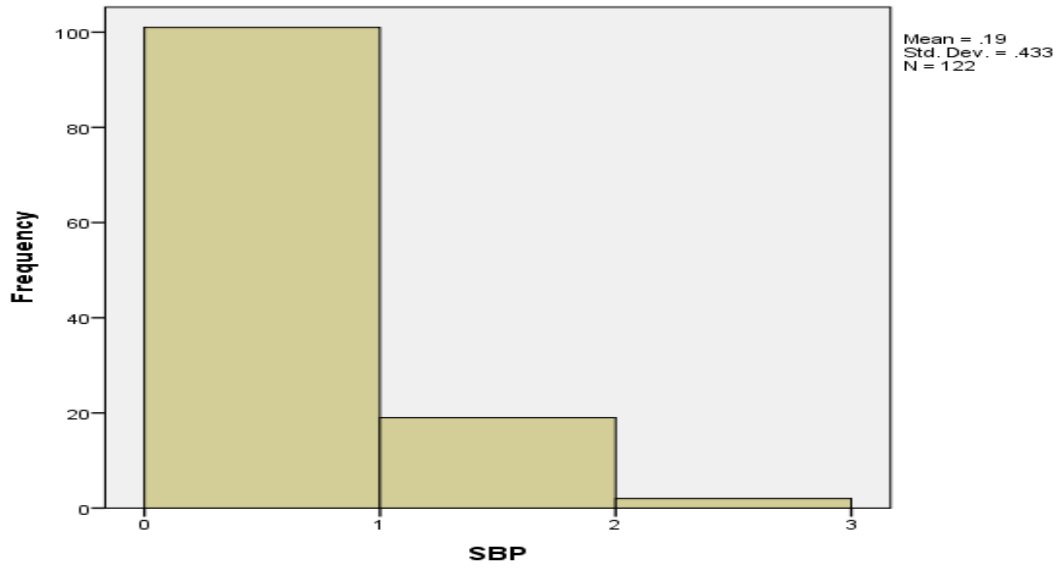


Figure 32 – Histogram representing UK1 'SBP' (Social Behavioural Positive) teacher comments (N=122)

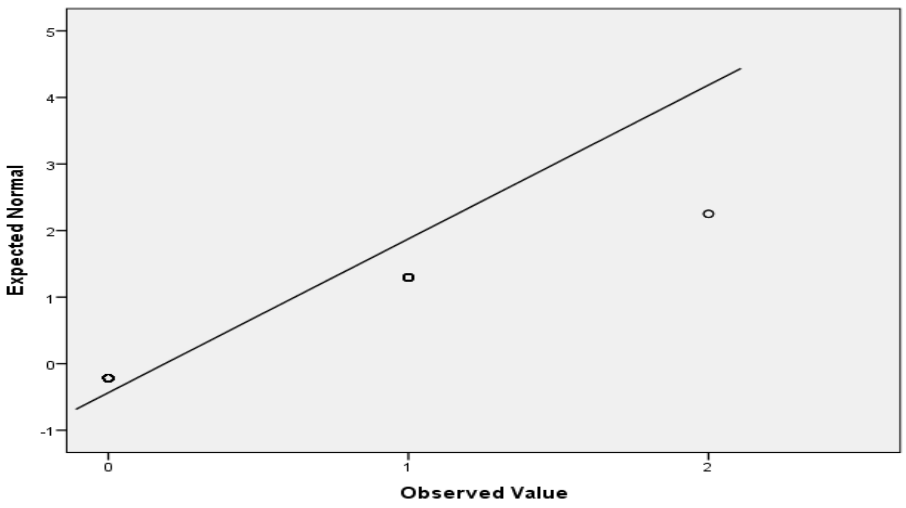


Figure 33 – Exponential Q-Q Plot of UK1 'SBP'

The exponential distribution of teachers' use of social behaviour positive comments 'SBP' (Figure 33 above), the small total number of 'SBP' comments in the UK1 dataset (mean= .19, N= 122) and the high number of zero scores (101 of 122) reflected the repeated theme in the literature (Wheldall *et al.,* 1989; see Chapter 2) that teachers in secondary schools do not usually congratulate students for good behaviour. This is similar to the distribution of 'SBP' scores in the UK27 dataset (mean= 2.89, N= 106).

5.10 As with 'SBP', the exponential distribution of teachers' use of task performance criticism 'TPC' comments, the relatively small total number of 'TPC' comments in the UK1 dataset (mean= .86, N= 122) and the high number of zero scores (69 of 122) reflected another repeated theme in the literature (Wheldall *et al.,* 1989; and see Chapter 2) and in the UK27 dataset (mean= 5.11, N=106) that teachers in secondary schools rarely criticise students for their work – at least in respect of making comments that the whole class and an observer would hear.

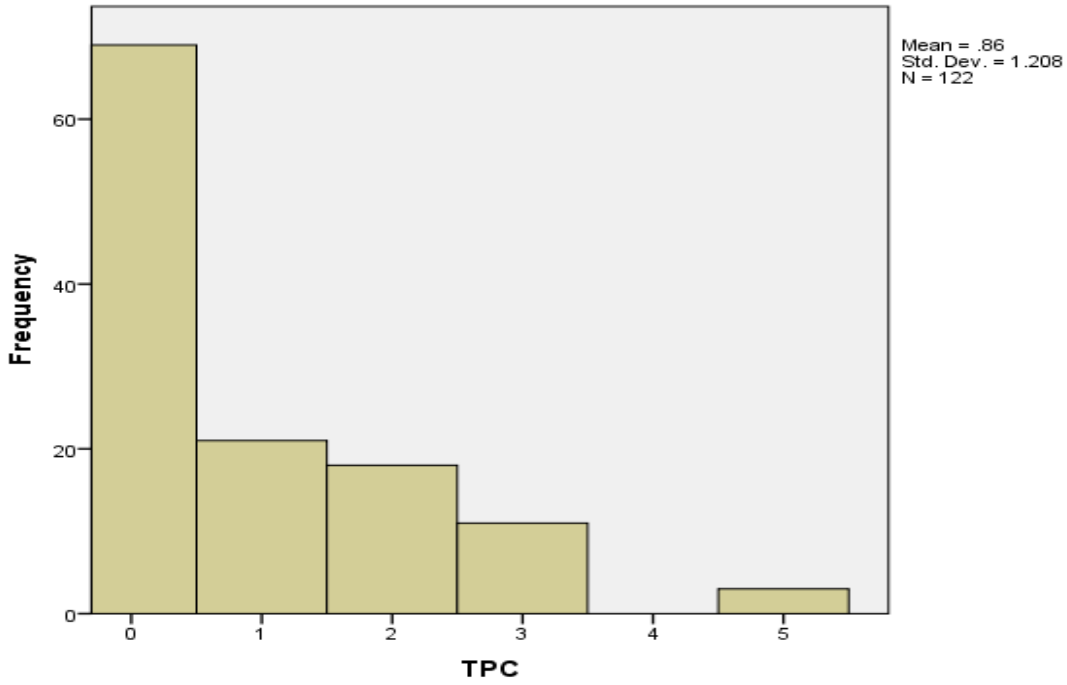


Figure 34 – Histogram representing UK 1 'TPC' (Task Performance Criticism) teacher comments (N= 122)

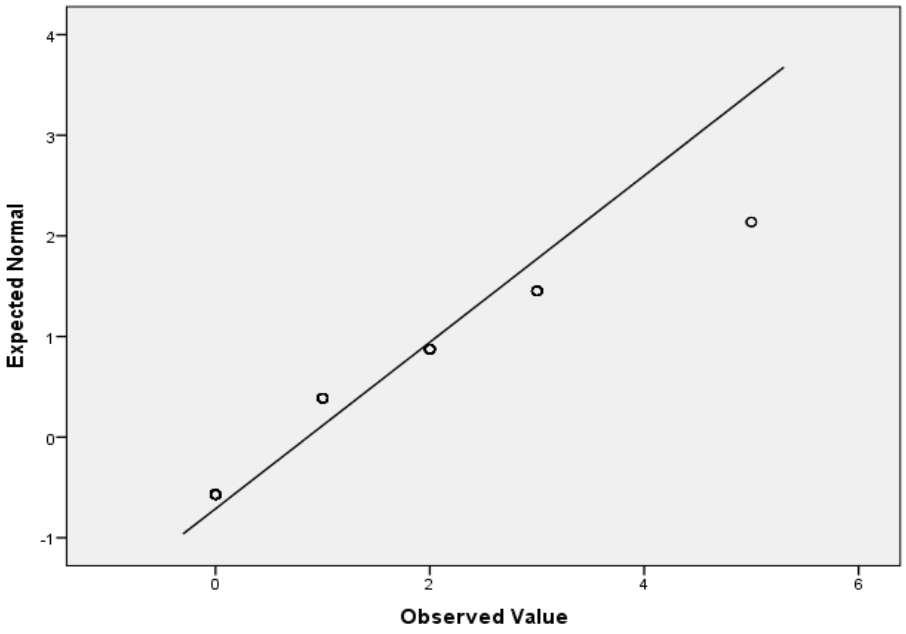


Figure 35 – Exponential Q-Q Plot of UK1 'TPC'

5.11 There appears to be a visible difference in the distribution of the UK1 and UK27 'SBC' distributions:

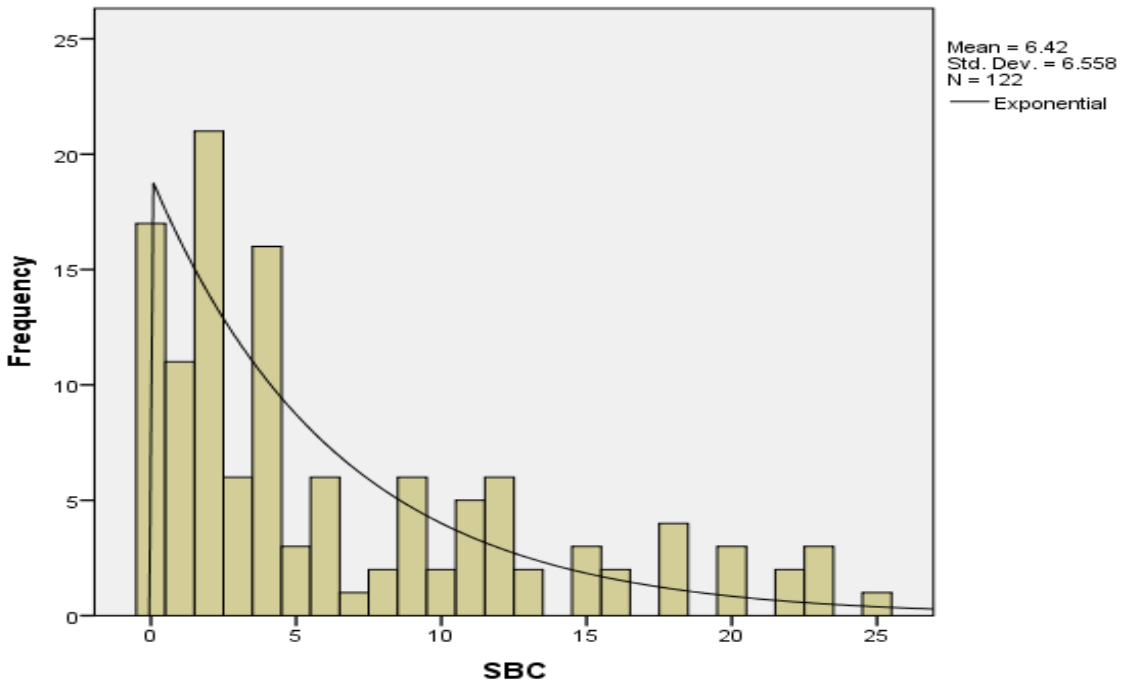


Figure 36 – Histogram representing UK1'SBC' (Social Behaviour Criticism) teacher comments with 'best-fit' exponential distribution curve (N=122)

The distribution is more exponential than normal and the mean much lower than for the UK27 ‘SBC’ data (mean= 12.21, SD= 9.87, N= 106).

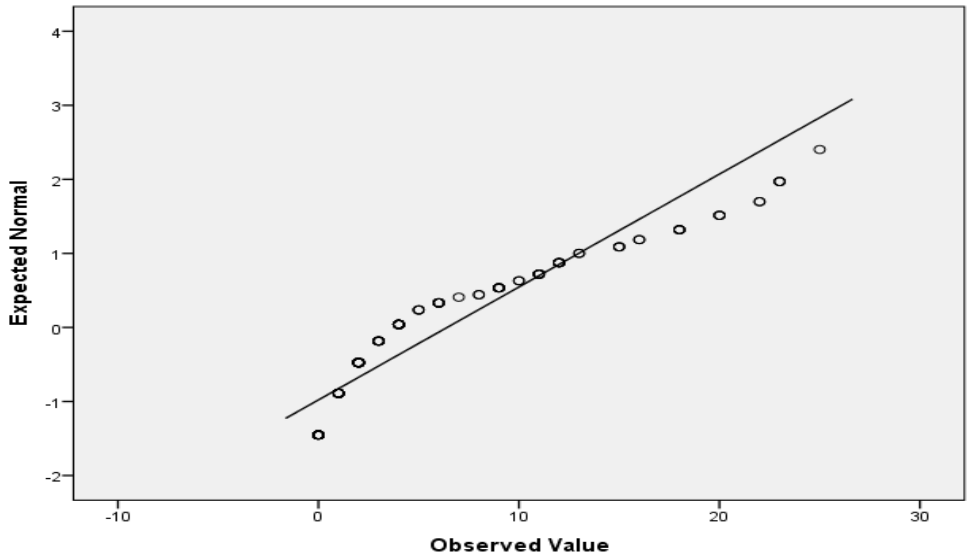


Figure 37 – Normal Q-Q Plot of UK1 ‘SBC’

5.12 An estimation of the mean amount that teachers spoke for in the UK1 dataset lesson observations - irrespective of their use of positive or negative feedback comments to students, was captured in the ‘INX%’ variable (neutral teacher-talk as a proportion of time as a percentage). The histogram and Q-Q Plot below depict a shallow left-skewed distribution:

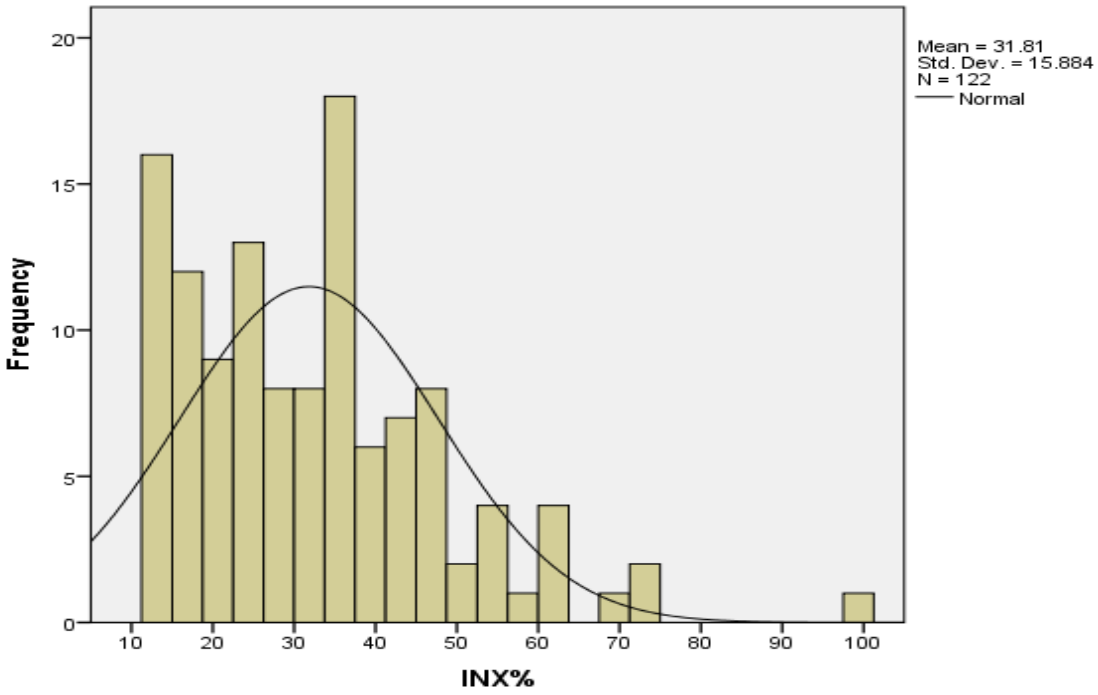


Figure 38 - Histogram representing UK1 'INX%' neutral teacher-talk as a proportion of time % (N=122)

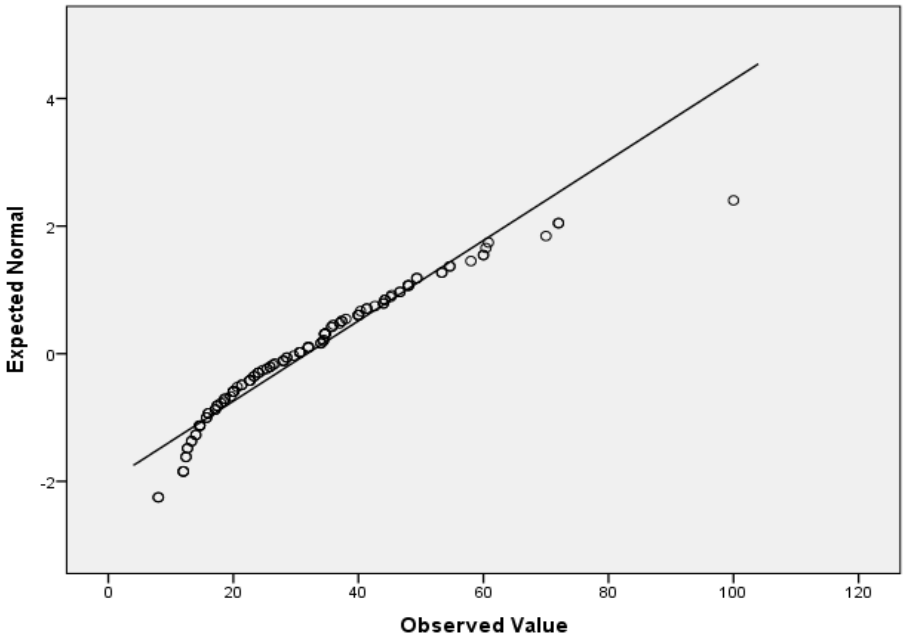


Figure 39 – Normal Q-Q Plot of UK1 'INX%'

Of note is the much lower UK1 mean 'INX%' figure (mean= 31.81, SD= 15.884, N=122) in comparison to the UK27 mean 'INX%' figure (mean= 52.66, SD= 18.91, N=106). Differences between the two datasets are tested for significance below (see Table 25).

5.13 Whilst 3 feedback types in the UK1 verbal feedback data violated the parametric requirement for normality, ('TPP'; 'SBP'; and 'TPC') with exponential distributions and high numbers of zero scores; the fourth feedback type ('SBC') with higher tallies, less zero scores and a mean of 6.42 presented a more marginal violation. The dependent variable ('On-task%') and the total measure of teacher-talk ('INX%') were marginally normal distributions.

As with the UK27 dataset, the statistical method of choice to identify the predictive relationship between a number of different dynamically independent variables, e.g. the amount of neutral teacher-talk: 'INX%'; and teacher-feedback comments ('TPP'; 'SBP'; 'TPC'; and 'SBC') and students following teacher-directions ('On-task%') was still parametric multiple regression (Tabachnick and Fidell, 2012) as colinearity between any of the five teacher-feedback independent-variables and the dependent variable 'On-task%' could be ruled out and there were no evident violations of the homoscedasticity assumption in scatter plots of distributions.

5.14 Hierarchical multiple regression analysis (HMRA) of UK1 teacher-feedback variables and 'On-task%'

Hierarchical multiple regression analysis (HMRA) was used to assess the relative predictive relationship between 5 control measures ('TPP'; 'SBP'; 'TPC'; 'SBC'; 'INX%') and the dependent variable: students following teacher-directions ('On-task%'). The initial procedure using SPSS indicated that 7 cases in the UK1 dataset (when N= 129) were causing violation of the procedure by exceeding the critical value of the Mahalanobis distance for 5 variables (k= 5) of 20.52 for n>100, $p < .001$ (Tabachnick and Fidell, 2012). By removing 7 identified outlier-cases from UK1, as Tabachnick and Fidell advise for HMRA, the maximum Mahalanobis distance decreased (k=5) from 92.52 to 21.184. The sample using 5 independent variables as control measures for HMRA was still sufficiently large (N=122) with up to a 5% reduction of the dataset being judged as acceptable (ibid.).

Whilst the means of the 5 variables being considered as control measures with HMRA were more valid than medians in each case, (Pett, 2016), the small size of means, large number of zero scores and exponential distributions (Tabachnick and Fidell, 2012) suggested caution about the inclusion of all 4 feedback types: 'TPP'; 'TPC'; 'SBP'; and 'SBC' in the HMRA but results subsequently revealed that *excepting* 'SBC', the influence of any single feedback type in any of the 6 models was negligible, probably because of the high number of zero

scores. ‘SBC’ was not, strictly speaking, an exponential distribution and whilst there were a high number of zero scores (17), the mode was 21 tallies of 2 ‘SBC’ feedback comments (N=122).

5.15 The hierarchical order of variable inclusion for HMRA of the predictive relationship of 5 UK1 independent teacher verbal-feedback variables when considering the distribution of the dependent variable ‘On-task%’ was: Model 1 – ‘SBC’; Model 2 - add ‘TPC’; Model 3 – add ‘SBP’ and ‘TPP’; Model 4 – add ‘INX%’; Model 5 – subtract ‘SBC’ and ‘TPC’; Model 6 – subtract ‘INX%’:

Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**
Single most predictive independent variable	*Critical teacher-feedback only*	*Praising and critical teacher-feedback*	*Comprehensive teacher verbal feedback model*	*All positive + teacher-feedback + neutral talk*	*Only praising feedback*
SBC	SBC	SBC	SBC	SBP	SBP
	TPC	TPC	TPC	TPP	TPP
		SBP	SBP	INX%	
		TPP	TPP		
			INX%		

Table 18 – Hierarchical multiple regression: order of models and UK1 teacher-feedback variables

5.16 ‘SBC’ accounted for 16.1% of the variance (F (1, 120) = 22.99, $p < .001$) of ‘On-task%’. ‘TPC’, when added in accounted for an additional .3% (F (2, 119) = 11.68, $p < .001$) meaning that critical teacher-feedback as measured by these variables accounted for 16.4% of the variance.

When added in, ‘SBP’ and ‘TPP’ accounted for an additional 3.1% (F (4, 117) = 7.11, $p < .001$) of the variance; and ‘INX%’ for only an additional .3% (F (5, 116) = 5.74, $p < .001$); so the total variance of ‘On-task%’ accounted for by the comprehensive teacher verbal feedback model - Model 4, that included the 4 teacher-feedback variables, ‘TPP’ (B= .173 p= .055); ‘SBP’ (B= .030 p= .730); ‘TPC’ (B= .040 p= .646); ‘SBC’ (B= -.464 p< .001) plus the measure of total observed teacher-talk, ‘INX%’ (B= .054 p= .527) was: 16.4%. Only the contribution of ‘SBC’ was individually significant (p< .005).

'SBC' and 'TPC' were subtracted to create Model 5 which consisted only of positive and / or neutral teacher-feedback: 'TPP'; 'SBP' and 'INX%'. Model 5 accounted for only 1.9% of the variance (F (3, 118) = .258, p=.856); and when 'INX%' was subtracted to create Model 6, the effect was negligible. Model 6 - 'TPP' and 'SBP' accounted for .3% of the variance (F (2, 119) = .178, p=.837) also.

The $\mathcal{B}$ (beta) statistic provided a relative measure for each predictor / control measure of their relative contribution to the comprehensive model; in order of importance within the UK1 dataset – the largest is listed first: 'SBC'; 'TPP'; 'TPC'; 'INX%'; and 'SBP'. This differs from the UK27 dataset in that 'INX%' and 'SBP' have swapped bottom places. With the UK1 dataset, as with the UK27 dataset, it is evident that 'SBC' provided by far the biggest single contribution in respect of accounting for the total variance in 'On-task%'. It is also important to note that the *direction* of contribution as indicated by a negatively-signed $\mathcal{B}$ in the cases of both 'SBC' and 'SBP' is *negative* - meaning that *less* of both is predictively associated with *more* 'On-task%'.

5.17 UK1 teacher-feedback and 'On-task%' correlations

As was expanded upon in respect of the UK27 dataset previously (4.21) the predictive relationship between any of the teacher verbal behaviour measures ('TPP'; 'TPC'; 'SBP'; 'SBC'; and 'INX%') and students following a teachers' directions (dependent variable 'On-task%') cannot be considered a causal one - in either direction (see Chapter 6 for a fuller discussion of these matters.) Correlational statistics are used here, as they were with the UK27 dataset, to provide supportive secondary information for associative relationships between variables and not as primary arguments about distribution predictability.

The next section investigates further the UK1 distributions of the four variables 'TPP; 'TPC'; 'SBP'; and 'INX%' and their less predictive relationship (as indicated by HMRA) with the distribution of the dependent variable 'On-task%' visually using scatter-plots and correlational statistics.

5.18 A positive correlation *expected* from the literature review and thus expressed *a priori* for the current research in Hypothesis #1 (see Chapter 1) was that 'On-task%' and 'TPP' would be significantly associated. There was no significant evidence that this was the case with the UK27 secondary school dataset from the parametric correlational analysis, and neither was there any evidence that this was the case with the UK1 dataset.

Investigating further the distributions of the variable UK1 'TPP' and its less predictive relationship with the distribution of the dependent variable 'On-task%' using scatter-plots, like the UK27 'TPP' distribution, the following UK1 'TPP' graph shows clustering of scores in the left-upper quadrant:

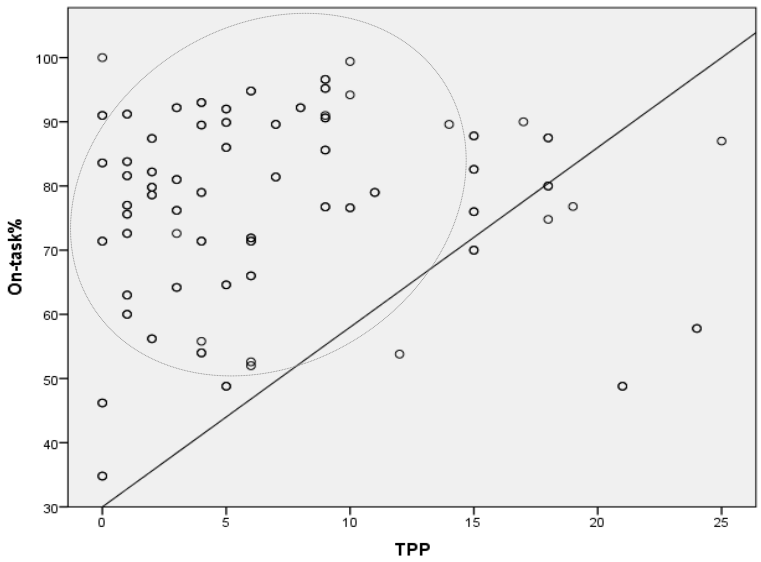


Figure 40 – Scatter plot showing UK1 'TPP' (Task Performance Positive) and its association with students following teacher-directions ('On-task%') (N= 122)

This distribution pattern is similar to UK27 'TPP' (see Figure 16) and to the pattern reported in the author's primary study (Apter, *et al.*, 2010, p. 156).

As reported above, the distribution of 'On-task%' was approximately normal but 'TPP' less so. Spearman's non-parametric rank correlation coefficient was insignificant (*rho*= .131, N= 122, *p*= .149, 2-tailed). This contrasted with the results from the author's primary school research (Apter, *et al.* 2010, p.155) where a *positive* correlation was calculated (*r*= .345, *p* <.01, two-tailed).

5.19 As noted above, the distribution of 'TPC' data points followed an exponential best-fit line with many tallies between 0 and 1 (see 5.10 above).

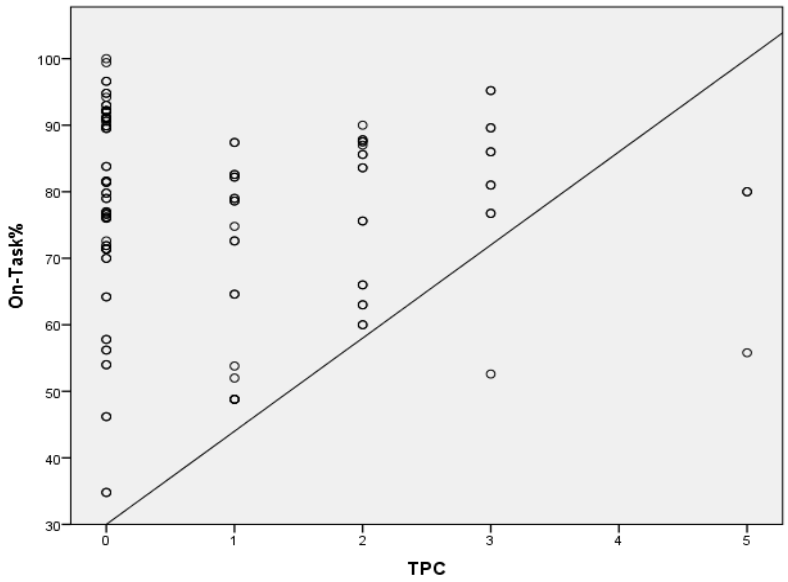


Figure 41 – Scatterplot showing UK1 'TPC' (Task Performance Criticism) teacher comments and the association with students following teacher-directions ('On-task%') (N=122)

The non-parametric correlation between the two variables using Spearman's product-moment correlation coefficient was insignificant (*rho*= -.075, N= 122, *p*= .414, two-tailed).

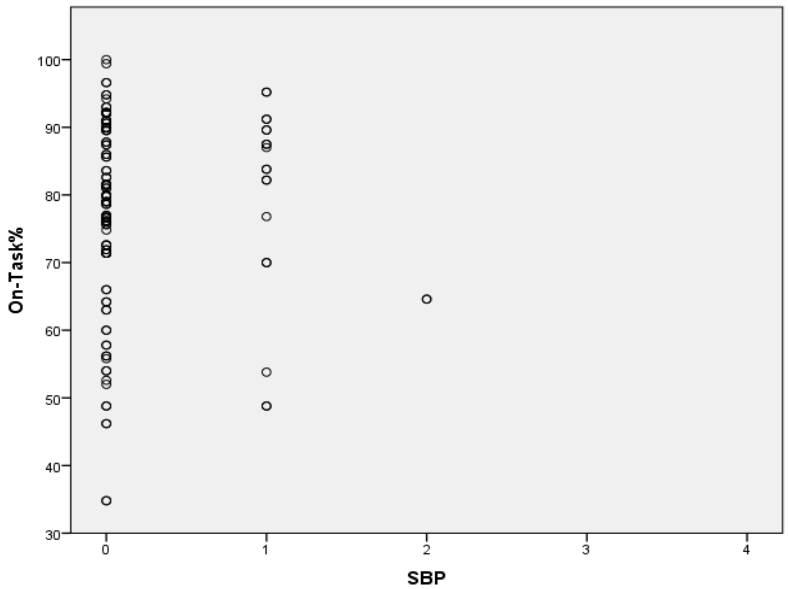


Figure 42 – Scatterplot showing UK1 'SBP' (Social Behaviour Positive) teacher comments and the association with students following teacher-directions ('On-task%') (N= 122)

In respect of 'TPC', the distribution of 'SBP' data points also approximated an exponential distribution with many tallies of 0 (mode= 101) and less tallies of 1 (median= 19) on the X

axis. The non-parametric correlation between the two variables using Spearman's rank-order correlation coefficient was insignificant (*rho*= .046, N= 106, *p*= .615, 2-tailed).

5.20 As noted in the previous chapter, another *positive* correlation *expected* from the literature review and thus expressed *a priori* for the current research in Hypothesis #2 (see Chapter 1) was that 'On-task%' and 'INX%' would be significantly associated. There was no significant evidence that this was the case with the UK27 secondary school dataset (see Chapter 4), either from the HMRA, or from a parametric correlational analysis (4.16). The UK1 HMRA indicated a similar outcome (see 5.16) and both[34] parametric and non-parametric correlations supported the null hypothesis (*r*= 0.60, N=122, *p*= .510; *rho*= 0.60, N=122, *p*= .509).

The following 'INX%' graph shows some clustering around the top-left of the chart:

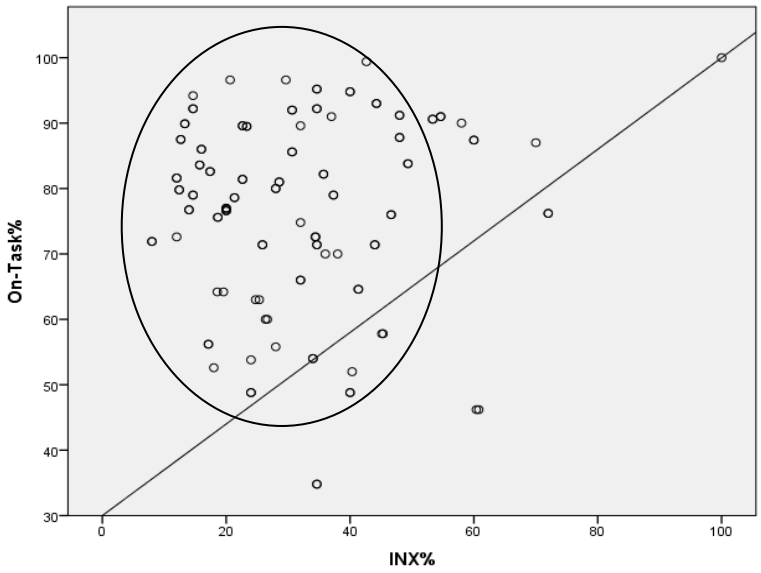


Figure 43 – Scatterplot showing UK1 'INX%' neutral teacher-talk and its association with students following teacher-directions ('On-task%') (N=122)

This suggests that teachers in the UK1 secondary school who talked for between 10% and 60% of their time in class were more likely to be teaching academic classes where students were following their directions, and that very few teachers in the UK1 secondary school talk more than 60% of the time in academic lessons.

[34] Given marginal normality of both distributions: 'On-task%' and 'INX%'.

5.21 UK1 teacher-feedback type ratios

In common with findings from earlier research, (Wheldall *et al.*, 1989; and see Chapter 2), UK27 teachers comments critical of students' behaviour 'SBC' were used more than 'SBP' and 'TPC' but not as frequently as 'TPP'. This section examines whether this holds true in the UK1 school and whether Hypothesis #3 (Chapter 1) was similarly supported.

The following ratio table provides means and a comparison of ratios of the 4 types of teacher-feedback comments used in the UK1 secondary school ('TPP'; 'SBP; 'TPC' and 'SBC'):

<table>
<tr><td>TPP Mean= 6.63
SD= 6.088
N = 122</td><td colspan="2">TPP / SBP 34.89 : 1</td><td>SBP Mean= .19
SD= .433
N = 122</td></tr>
<tr><td>TPP / TPC
7.71 : 1</td><td>TPP / SBC
1.03 : 1</td><td>TPC / SBP
4.53 : 1</td><td>SBP / SBC
1 : 33.79</td></tr>
<tr><td>TPC Mean= .86
SD= 1.208
N = 122</td><td colspan="2">TPC / SBC 1 : 7.46</td><td>SBC Mean= 6.42
SD= 6.558
N = 122</td></tr>
</table>

Table 19 - Ratio table of UK1 teacher-feedback means: 'TPP'; 'SBP; 'TPC'; and 'SBC'

A one-way repeated measures ANOVA was conducted to compare the means of UK1 teacher-feedback comments in 4 categories: 'TPP'; 'SBP; 'TPC'; and 'SBC'. There was a significant effect for teacher-feedback comments, Wilk's lambda, $\Lambda = .31$, (F (3, 114) = 85.19, $p < .005$), multivariate partial *eta* squared = .69 which is characterised as a large effect size (Tabachnick and Fidell, 2012). Pairwise comparisons revealed that excepting a comparison between 'TPP' and 'SBC' where the difference is insignificant (and irrelevant to Hypothesis #3), each of the 4 category tallies are significantly different from each of the 3 other category tallies, $p < .05$.

Hypothesis #3 is therefore supported by the UK1 dataset: teachers in the UK1 secondary school use significantly more *positive* verbal feedback, e.g. praising comments, to *encourage* academic-task-focussed behaviour and product; and significantly more *negative* verbal feedback, e.g. critical comments, to discourage undesired social-behaviour ('Off-

task' behaviour); and use significantly less *positive* verbal feedback to encourage students to follow their social-behavioural (non-academic) directions; and significantly less *negative* verbal feedback to discourage undesired, incorrect or inaccurate academic-task-focussed behaviour or product.

5.22 UK1 teacher-feedback types by year-group

For the UK1 dataset, a *post hoc* Bonferroni procedure was used to test the significance of differences between year groups in respect of the use made by teachers of the 4 types of teacher-feedback. Treated as a single phenomenon, there were evident differences between the way that feedback was used in Year 11 and Year 12, and the lower years: 7 to 10 and less feedback was used with older students. Particularly, all feedback types were used significantly less ($p < .005$) in Years 11 and 12 in comparison to Years 7 and 9; but not 8 and 10:

(I) 'Year group'	(J) 'Year group'	(I-J) Mean Difference	Standard Error	Significance
Y11	Y7	-2.83*	.660	.001
	Y8	-.81	.690	1.0
	Y9	-2.44*	.690	.009
	Y10	-1.14	.745	1.0
	Y12	.95	.729	1.0
Y12	Y7	-3.77*	.778	.000
	Y8	-1.76	.803	.463
	Y9	-3.39*	.803	.001
	Y10	-2.09	.851	.236
	Y11	-.95	.729	1.000

*Mean difference is significant at the .05 level

Table 20 – Mean differences of 4 combined teacher-feedback type ('TPP'; 'TPC'; 'SBP'; 'SBC') compared by 'Year group' calculated using Bonferroni *post hoc* procedure

5.23 UK1 teacher-feedback additional findings

An *expected* correlation (following the results from the UK1 HMRA reported above) more fully revealed by a statistical exercise using Spearman's rank-order correlation coefficient

was the finding of a negative medium correlation between 'SBC' and 'On-task%' (*rho*= -.431, N= 122, significant at the .01 level, 2-tailed).

This confirmed the implication of the UK1 HMRA and of the UK27 HMRA (see 4.22) that academic classes in the UK1 secondary school that were generally well-behaved, where students were mostly following teacher-directions ('On-task%'), were less likely to provoke social behaviour criticism ('SBC') from the teacher.

5.24 As with the UK27 dataset, in the UK1dataset, 'SBC' was significantly *positively* correlated with 'TPP' (*rho*= .240, *p*< .01, 2-tailed); and with 'TPC' (*rho*= .232, p< .05, 2-tailed) meaning that teachers in the UK1 secondary school who used critical social behaviour feedback were also more likely to use 2 other forms of verbal feedback: praising and critical comments directed towards students' academic behaviour. A similar relationship was identified between 'SBP' and 'TPP' (*rho*= .202, *p*< .05, 2-tailed).

5.25 In respect of the total amount of talking a teacher used in the course of teaching an academic lesson ('INX%'), there were no significant correlations with 3 of the 4 UK1 teacher verbal feedback types ('TPP'; 'TPC'; 'SBC') and a weak correlation between 'SBP' and 'INX%' (*rho*= .182, *p*< .05, 2-tailed).

Findings in respect of the UK1 and UK27 datasets were similar; neither 'TPP' nor 'INX%' were correlated with 'On-task%' and therefore contrast with the author's findings in primary schools. This was a surprising finding and requires further investigation. In the primary study both task-performance positive and the total amount of teacher talk were significantly correlated with on-task behaviour (see 4.24; Apter, *et al.,* 2010, p.155).

5.26 UK1 Contextual variables descriptives

HMRA was also used to assess the relative predictive relationship between 6 UK1 school contextual factors: 'Years' experience'; 'Teacher gender'; 'Year group'; 'Class size'; 'Adults present'; and 'AM or PM'; and the dependent variable: 'On-task%' (students following teacher-directions). The 7th independent variable 'Subject' was excluded because within the UK1 dataset it was a non-dichotomous categorical variable with 6 levels (academic subjects): English, maths, science, history, geography, and other (including media, modern languages and personal development). As with the UK27 dataset, 'Year group' was treated as a continuous variable for the purpose of HMRA (see 4.25).

5.27 The maximum Mahalanobis distance for 6 variables (*k*= 6) was 16.81 for n>100, *p* <.05 (Tabachnick and Fidell, 2012) and no cases in the data exceeded this. The maximum Mahalanobis distance of the UK1 data was 15.538. The means of the 6 variables being considered as control measures with HMRA were more valid than medians in each case, (Pett, 2016). 'Teacher gender' (scores 1 or 2, mean= 1.57, SD= .50); and 'AM or PM' (scores 1 or 2, mean= 1.48, SD= .502) were dichotomous variables and therefore admissible for HMRA despite violating normality (ibid.). Distributions of the 6 UK1 contextual variables failed the Kolmogorov-Smirnov test of normality but this was not judged a fatal violation of the assumption of normality (Palint, 2010).

5.28 The 'Year group' distribution had a left-skewed curve which could be described as a self-selection effect (Ziliak and McCloskey, 2008) with a mean of 9.51 and a 5% trimmed mean of 9.51. Skewness was measured at *S*= -.103 (SE= .219); and kurtosis at *K*= -1.353 (SE= .435):

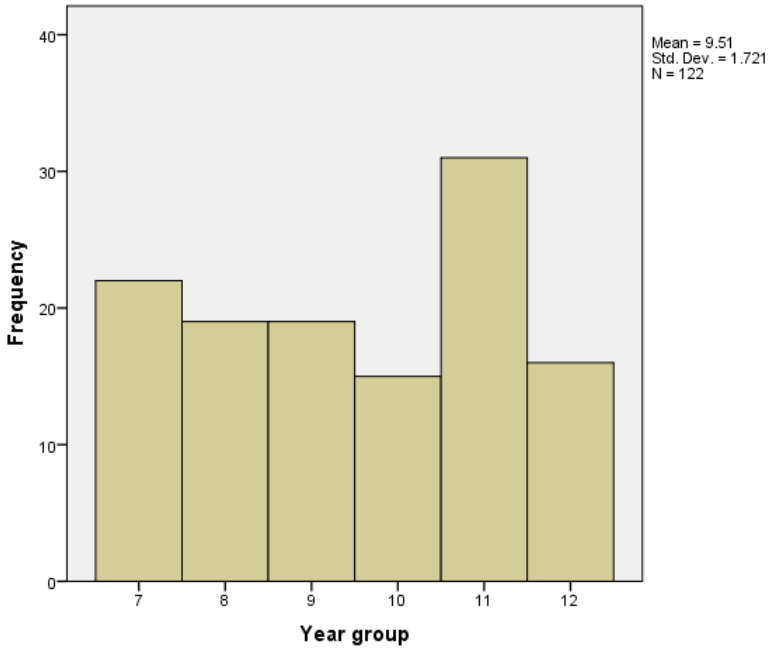


Figure 44 – Histogram showing distribution of UK1 'Year group' frequencies (N=122)

As can be seen from this histogram, Year 11 and Year 7 classes were selected most frequently by observers for the current research because the UK1 school's senior leadership tam (SLT) was particularly interested in the results that the author obtained from these year groups.

5.29 The distribution of the UK1 'Years' experience' frequencies accords to a left-skewed high peaked normal distribution with a mean of 10.31 years (compared to the UK27 figure of 8.2 years) and a 5% trimmed mean of 9.66. Skewness was measured as *S*= -.302 (SE= .219); and kurtosis as *K*= -1.941 (SE= .435):

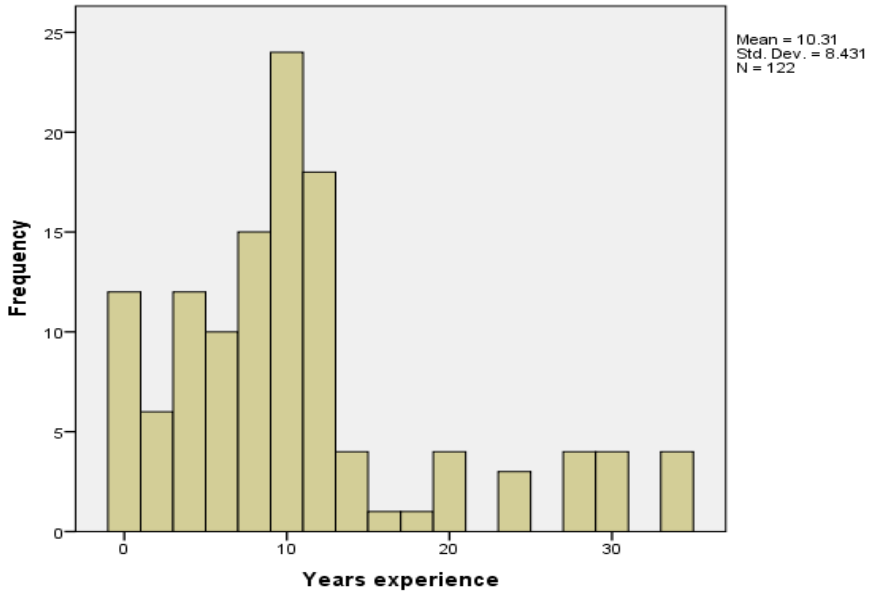


Figure 45 – Histogram showing distribution of UK1 'Years' experience' frequencies (N=122)

This distribution suggests that teachers stay teaching in the UK1 school on average 2 years longer than the UK average (DfE, 2015b).

5.30 As with the UK27 dataset, 'Adults present' in class was an exponentially shaped distribution for obvious reasons:

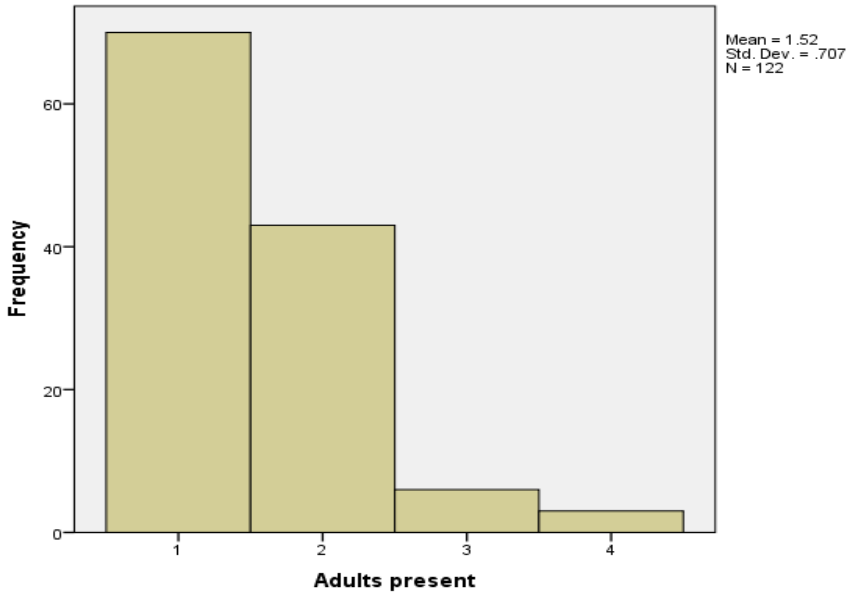


Figure 46 – Histogram of UK1 'Adults present' in class (N= 122)

5.31 The distribution pattern of UK1 'Class size' variable is suggestive of a normal distribution but the spike: 'Class size'= 26, is difficult to explain. The UK1 school's Senior Leadership Team were unable to provide further explanation. The UK1 'Class size' mean at 17.19 is 2 students smaller than the UK27 mean at 19.43.

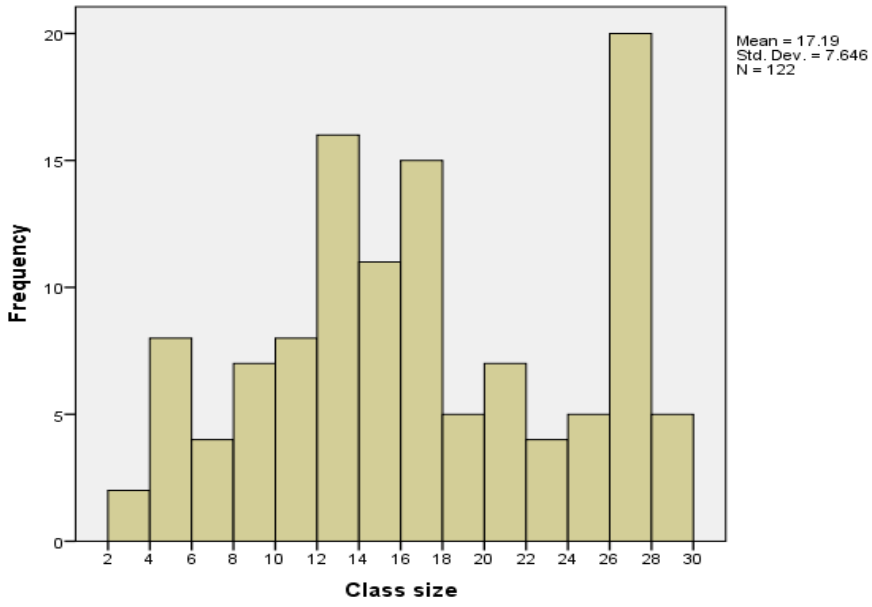


Figure 47 – Histogram of UK1 'Class size' (N=122)

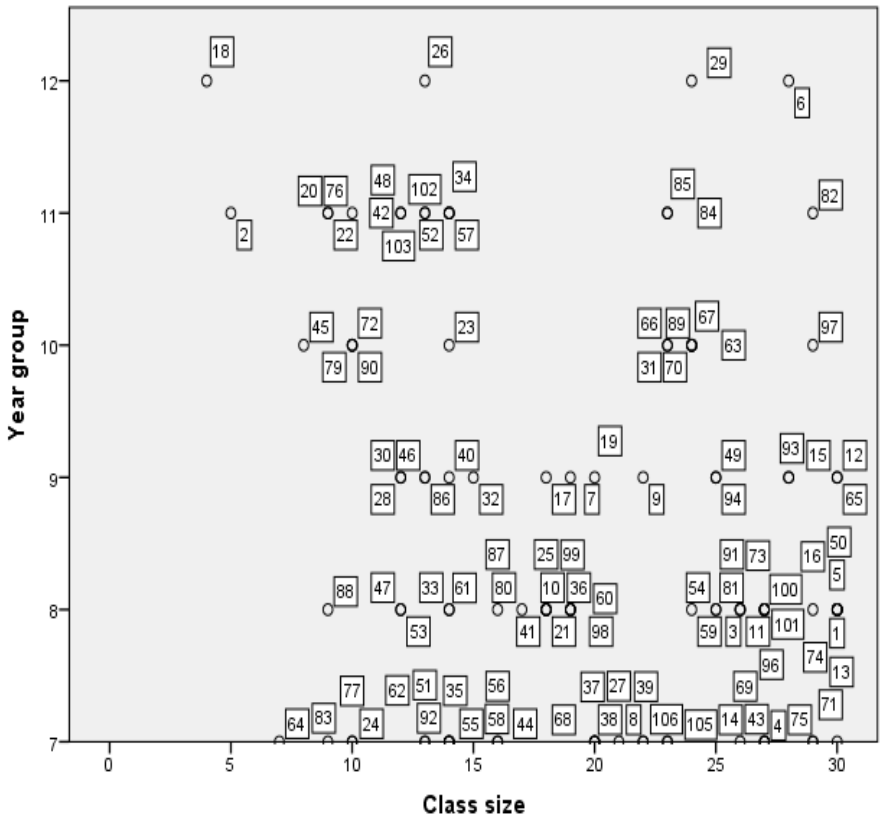


Figure 48 – Case-numbered scatter plot of UK1 'Class size' by 'Year group' data-points (N= 122)

The clustering evident in this UK1 scatter plot is remarkably similar to that noted with the UK27 plot (see 4.30).

5.32 As with the UK27 dataset, UK1 'Subjects' is a category variable, the levels (or categories) of which are separate curricular subjects (see also 4.31). 'Subjects' will not therefore be included in the UK1 HMRA of contextual variables but will be separately analysed. The histogram provides a representation of the distribution of curricular subject-lessons that were observed and included in the current research:

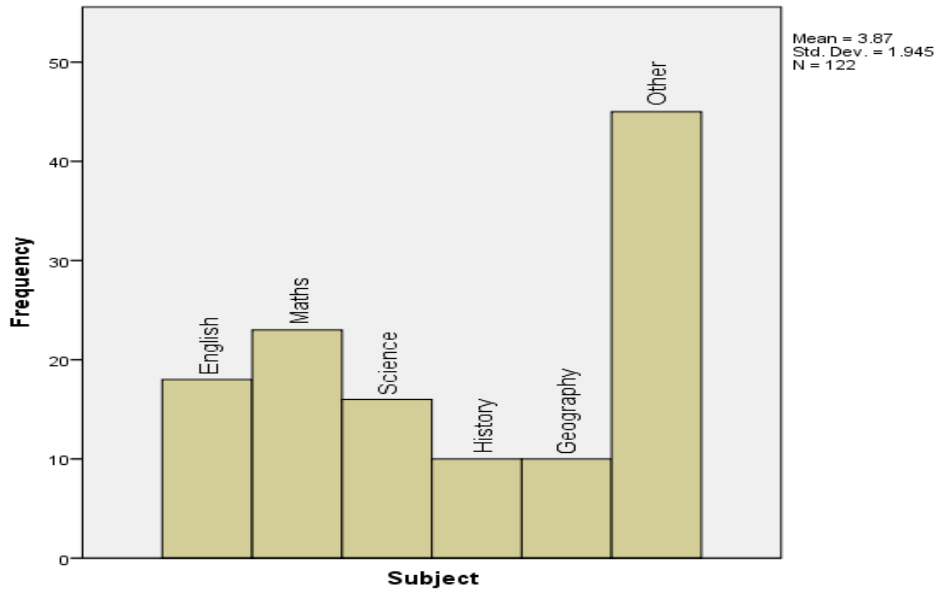


Figure 49 – Histogram showing distribution of curricular UK1 'Subjects' observed-lessons frequencies (N= 122)

5.33 UK1 HMRA of contextual variables

The hierarchical order of variable inclusion for HMRA of the predictive relationship of 6 UK1 independent contextual variables when considering the distribution of the dependent variable 'On-task%' was: Model 1 – 'Teacher gender' and 'Years' experience'; Model 2 - add 'Year group'; Model 3 – add 'Class size' and 'Adults present'; Model 4 – add 'AM or PM':

Model 1	Model 2	Model 3	Model 4
Two most predictive teacher-specific contextual variables	*Hypothetically next most predictive contextual variable*	*Add human context: class size plus no. of adults*	*All positive + teacher-feedback + neutral talk Only praising feedback*
Teacher gender Years' experience	Teacher gender Years' experience	Teacher gender Years' experience	Teacher gender Years' experience
	Year group	Year group	Year group
		Class size Adults present	Class size Adults present
			AM or PM

Table 21 – Hierarchical multiple regression: order of introduction of models and UK1 contextual variables

5.34 Consisting of two teacher-specific contextual factors: 'Teacher gender' and 'Years' experience', Model 1 accounted for 8.2 % of the variance (F (2, 119) = 5.34, $p < .01$) in 'On-task %'.

Model 2 - when 'Year group' was added in to the HMRA, accounted for an additional 11.5% of the variance (F (3, 118) = 9.68, $p < .001$). Model 3, when 'Class size' and 'Adults present' were added in, accounted for an additional 5% of the variance (F (5,116) = 7.62, $p < .001$). Model 4 - when 'AM or PM' was added in to the HMRA, accounted for an additional .4% of the variance (F (6, 115) = 6.42, $p < .001$).

Model 4 - the comprehensive contextual model which included 6 UK1 contextual variables (excepting the inclusion of 'Subject'), accounted for a total 25.1% of the variance in the 'On task %' distribution. Model 4 consisted of (in order of predictive importance – the largest $\mathcal{B}$ first): 'Class size' ($\mathcal{B}$= -.250, $p < .05$); 'Year group' ($\mathcal{B}$= .238 $p < .05$); 'Years' experience'($\mathcal{B}$= .221 $p < .05$); 'Teacher gender' ($\mathcal{B}$= -.142 p= .092)[35]; 'AM or PM' ($\mathcal{B}$= .06, p= .458); 'Adults present' ($\mathcal{B}$= .039 p= .639). The contributions of 'Class size'; 'Year group'; and 'Years' experience' were individually significant ($p < .05$).

It is important to note that 'Class size' appeared to be a more important and predictive variable in the UK1 dataset than analysis revealed it to be in the UK27 dataset (Chapter 4). This difference was emphasised by a significant negative correlation using Spearman's rank-order coefficient made in the analysis of the UK1 dataset (*rho*= -.333, N=122, p< .001; above) made between 'Class size' and 'On-task%' – meaning students appeared to be more likely to be following teachers' directions when the class-size was smaller in the UK1 school[36].

5.35 Other correlations in the UK1 dataset involving contextual variables

There were a number of emergent un-hypothesised contextual-variable correlations in the UK1 dataset revealed by using Spearman's rank order coefficient, *rho*. As these correlations were not hypothesised and only a small number have relevance for the research questions of the current research (Chapter 2), the annotation is brief.

[35] Negatively signed statistic relates to dichotomous variable, male= 1, female= 2; therefore, ($\mathcal{B}$= -.189 $p < .06$) indicates that male teachers are more predictively associated with higher 'On-task %' but at an insignificant level ($p > .05$).

[36] There was no significant correlation between 'Class size' and 'On-task%' in the UK27 dataset.

In the UK1 dataset, 'Teacher gender' was correlated with 'Year group' (*rho*= -.217, N= 122, *p*< .05); 'Years teaching' with 'Year group' (*rho*= .375, N= 122, *p*< .01) and 'Teacher gender' (*rho*= -.264, N= 122, *p*< .01); 'Subject' with 'Teacher gender' (*rho*= -.320, N= 122, *p*< .01) and 'Years teaching' (*rho*= .274, N= 122, *p*< .01); 'Adults present' with 'Years teaching' *rho*= -.313N= 122, *p*< .01) and 'Subject' (*rho*= -.272= 122, *p*< .01); 'Class size' with 'Year group' (*rho*= -.421= 122, *p*< .01) and 'Subject' (*rho*= -.216122, *p*< .05).

It would appear from these correlations that in the UK1 school, male teachers are more likely to be teaching the upper years, have more years teaching experience and be teaching smaller classes.

5.36 UK1 teacher-feedback by year-group

Looking provisionally at the relationship between 'Year group', 'On-task%' and the 4 teacher verbal feedback types: 'TPP'; 'TPC'; 'SBP'; and 'SBC'; and the neutral teacher verbal behaviour percentage, 'INX%', it is instructive to use a line-graph to visually depict trends in the UK1 secondary school, year-group by year-group:

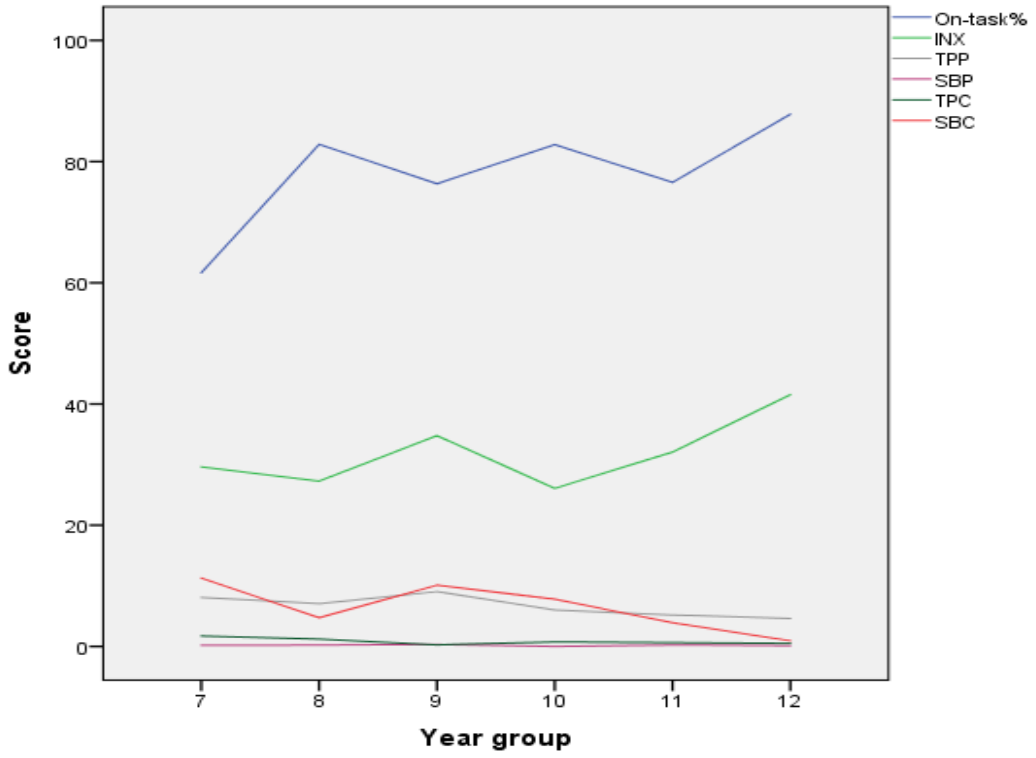


Figure 50 – line graph depicting the UK1 frequency of 6 variables – 'On-task%' and 5 teacher verbal behaviour variables by 'Year group'

Comparing the upward trend in the UK1 'On-task%' means from left to right with the gentle downward trend evident in the UK27 dataset (see 4.35) suggests that the baseline from which students start in Year 7 immediately following their transfer from primary school is influential. In the UK27 dataset the starting-point in Year 7 is students following a teachers' directions for 82% ('On-task%') of the time in academic lessons; and the end-point in Year 12 is 74% (-8%).

With the UK1 school's dataset students begin in Year 7 at a much lower 62% 'On-task%' figure but improve as they move up the school to an 88% 'On-task%' figure (+26%). The graph suggests that students and teachers in Year 7 in the UK1 secondary school increase the 'On-task%' by 20% during the students first year with them. This seems to be a remarkable achievement.

The graph also depicts how teachers use more 'TPP' and 'SBC' than 'SBP' and 'TPC' in each year group but that the tally-means converge downwards to low means in Year 11 and 12. This reflects maybe the expectation that older students are able to manage with less teacher support including verbal feedback.

Emulating the procedure used with the UK27 dataset (see 4.35), the following table displays 'On-task%'; and the 2 most predictive teacher-feedback types, 'TPP' and 'SBC' categorised by UK1 school's year groups:

	Year group	Mean	SD	N
On-task%	7	61.65	15.927	22
	8	82.84	10.891	19
	9	76.35	14.028	19
	10	82.80	8.026	15
	11	76.57	11.519	31
	12	87.81	8.667	16
	Total	77.06	14.442	122
TPP	7	8.09	7.702	22
	8	7.05	5.137	19
	9	9.05	6.553	19
	10	6.00	5.014	15
	11	5.19	6.025	31
	12	4.63	4.113	16
	Total	6.63	6.088	122
SBC	7	11.27	6.033	22
	8	4.74	4.382	19
	9	10.11	7.256	19
	10	7.80	8.082	15
	11	3.90	4.721	31
	12	.94	1.340	16
	Total	6.42	6.558	122

Table 22 – Means and standard deviations by 'Year group' of 'On-task%'; 'TPP'; and 'SBC' in the UK1 dataset

The data presented in this table emphasises the upward trend of UK1 'On-task%' and the reciprocating downward trends of 'TPP' and 'SBC'.

5.37 A one-way between-groups multivariate analysis of variance (MANOVA) was performed on the UK1 dataset to investigate 'Year group' differences upon classroom behaviour – both teacher and student. As with a similar analysis of the UK27 dataset (see 4.36), 'TPP' and 'SBC' were added-in with 'On-task%' as combined dependent variables to provide a fuller picture of the interdependence and covariance of teachers' verbal feedback with students' 'On-task%' behaviour.

The independent variable was 'Year-group' which for the purpose of MANOVA was now treated as a categorical variable[37] with 5 groupings (Y7, Y8, Y9, Y10, Y11 and Y12). Preliminary assumption testing was conducted to check for normality, linearity, univariate and multivariate outliers, homogeneity of variance-covariance matrices, and multicollinearity.

The 3 dependent variables grouped together failed Box's test; and the 3 dependent variables 'On-task%'; 'TPP' and 'SBC' initially failed Levene's test of equality of error variances. Given the unequal and small number of 'TPP' and 'SBC' data in each cell (< 25), the Box's Test failure (an upheld null hypothesis) was anticipated (Tabachnick and Fidell, 2012) and the test was not re-run, but the 3 variables were transformed arithmetically[38] and Levene's test run again. All 3 transformed variables passed the Levene test ('On-task%': $F(5, 96)= 2.12$, $p> .05$; 'TPP': $F(5, 96)= 1.40$, $p> .05$; and 'SBC': $F(5, 96)= 2.29$, $p> 05$).

There were statistically significant differences between 'Year group' (year-groups 7 to 12) on the combined dependent variables: 'On-task%'; 'TPP' and 'SBC', $F(15, 259.89)= 5.31$, $p= 0.18$; Wilks' *lambda*= .478; partial *eta* squared= 0.218 – a *large* effect size. When the results for the dependent variables were considered separately, both 'On-task%' and 'SBC' were significantly different when measured comparatively by 'Year group', with a *large* effect size in each case (Bonferroni-adjusted *alpha* level of 0.17; On-task%': $F(5, 96)= 7.31$, partial *eta* squared= .276; 'SBC': $F(5, 96)= 10.99$, partial *eta* squared= .364). The difference in 'TPP' by 'Year group' did not reach statistical significance ($p< .017$).

Pertinently, but perhaps unsurprisingly, teachers in the UK1 secondary school used more 'SBC' feedback - criticism of social behaviour in the lower years and less in the upper years.

[37] For HMRA, 'Year group' had been treated as a continuous variable.
[38] 'On-task%2'=SQRT (K-On-task%); 'TPP2'=LG10(TPP); 'SBC2'=LG10(SBC)

5.38 UK1 teacher-feedback by subject

Whilst 'Subject' could not be included in the HMRA of UK1 contextual variables because of it being a categorical variable, the correlations between curricular 'Subject' with the dependent variable 'On-task%' can be reported. The relationship between 'Subject' dummy-variables, 'English', 'Maths' and 'Science' and the 'On-task%' dependent variable using the non-parametric (*rho*) correlational analysis revealed a significant negative relationship between 'Maths' and 'On-task%' behaviour (*rho*= -.259, *p*< .005) – meaning that students were far more likely to be observed *not* following teachers' directions in maths lessons in the UK1 secondary school[39]. There was no such association revealed by a similar analysis of the UK27 dataset. This suggested that the relationships between 'Subject' and 'On-task%' should be investigated further within the UK1 dataset:

	Subject	Mean	SD	N
On-task%	English	81.99	13.347	18
	Maths	71.87	10.609	23
	Science	72.26	16.987	16
	History	75.46	11.343	10
	Geography	70.16	23.258	10
	Other	81.34	12.271	45
	Total	77.06	14.442	122
TPP	English	7.72	5.245	18
	Maths	7.35	7.408	23
	Science	3.88	5.749	16
	History	14.30	6.550	10
	Geography	3.60	3.026	10
	Other	5.78	4.714	45
	Total	6.63	6.088	122
SBC	English	6.17	7.318	18
	Maths	9.17	7.408	23
	Science	6.62	6.228	16
	History	9.90	8.517	10
	Geography	6.40	2.875	10
	Other	4.27	5.374	45
	Total	6.42	6.558	122

Table 23 – Means and standard deviations by 'Subject' of 'On-task%'; 'TPP'; and 'SBC' in the UK1 dataset

[39] The Senior Leadership Team of UK1 Secondary school were particularly interested in the way that maths lessons were represented in this research and any comparison could be made with English lessons. Ofsted had been very concerned about maths lessons and outcomes when they had put the school into 'special measures' immediately prior to the observations that are reported here.

The table shows that the mean 'On-task%' rate in English lesson in the UK1 secondary school is much higher (+10% approximately) than maths and all other subject areas excepting 'Other.' The contrast between maths and English 'On-task%' rates, by year-group is further illustrated in the following line graph:

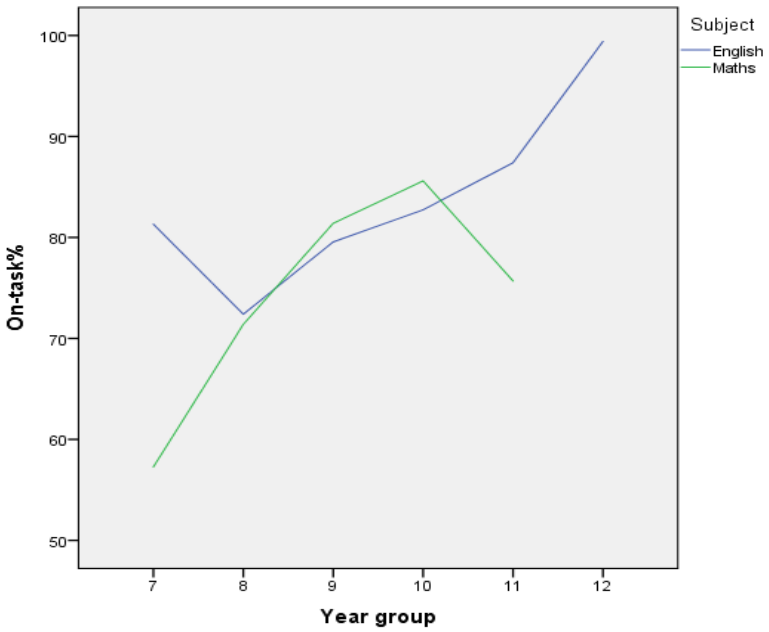


Figure 51 – A line graph of 'On-task%' rates by 'Year group' in UK1 English and maths lessons

In contrast to the UK27 dataset there was evidence that UK1 teachers used significantly different patterns or amounts of verbal feedback and neutral verbal behaviour ('TPP'; 'TPC'; 'SBP'; 'SBC'; 'INX%') between maths and English.

A paired-samples t-test was conducted to evaluate the impact of the two lesson-subjects, maths and English, upon 'On-task%'; 'TPP'; 'TPC'; 'SBP'; 'SBC' and 'INX%'. There was a statistically significant *decrease* in students following teachers' directions in maths ('On-task%': *M*= 71.87%, SD= 10.61) in comparison to English (*M*= 81.99%, SD= 13.35), t (17) = 2.13, $p < .05$. The mean decrease in 'On-task%' was 9.84% with a 95% confidence interval ranging from .075 to 19.60. The *eta* squared statistic (.21) indicated a large effect size (Cohen, 1988).

There was also a small but significant increase in UK1 teachers' use of critical verbal feedback directed at academic work ('TPC') in maths (*M*= 1.13, SD= 1.10) in comparison to English (*M*= .56, SD= .86), t (17) = -3.29, $p < .05$. The mean increase in 'TPC' was .57

with a 95% confidence interval ranging from -1.28 to -.28. There were no other significant differences.

5.39 A one-way between-groups multivariate analysis of variance (MANOVA) was performed on the UK1 dataset to investigate 'Subject' differences upon classroom behaviour – both teacher and student. As with the first analysis using MANOVA with the UK1 dataset and the MANOVA used with the UK27 dataset , 'TPP' and 'SBC' were added-in with 'On-task%' as combined dependent variables to provide a fuller picture of the interdependence and covariance of teachers' verbal feedback with students' 'On-task%' behaviour.

The independent variable was 'Subject', a categorical variable with 6 groupings (English, maths, science, history, geography, and other – see 5.02). Preliminary assumption testing was conducted to check for normality, linearity, univariate and multivariate outliers, homogeneity of variance-covariance matrices, and multicollinearity.

The 3 dependent variables 'On-task%'; 'TPP' and 'SBC', transformed[40] and grouped together, passed Box's test, (p>.001). Of the 3, 2 variables passed the Levene test ('TPP': F(5, 96)= 2.38, p> .04; 'SBC': F(5, 96)= 1.55, p> 05). 'On-task%': F(5, 96)= 4.072, p= .002) failed the test.

There were statistically significant differences between 'Subject' (English, maths, science, history, geography, and other) on the combined dependent variables: 'On-task%'; 'TPP' and 'SBC', F(15, 259.89)= 5.31, p< .005; Wilks' *lambda*= .703; partial *eta* squared= 0.111 – a medium effect size. When the results for the dependent variables were considered separately, only 'TPP' was significantly different when measured comparatively by 'Year group', with a moderate effect size (Bonferroni-adjusted *alpha* level of 0.17; 'TPP': F(5, 96)= 4.195, partial *eta* squared= .179). The difference in 'On-task%' and 'SBC' by 'Subject' did not reach statistical significance (p< .017).

Pertinently, but perhaps unsurprisingly, teachers in the UK1 secondary school used more 'TPP' feedback – positive comments about academic work in English lessons and less in maths lessons.

[40] 'On-task%'=SQRT (K-On-task%); 'TPP'=LG10(TPP); 'SBC'=LG10(SBC)

5.40 For the UK1 secondary school dataset, an additional *post hoc* finding was that teachers in the UK1 secondary school use less positive feedback to older students in academic lessons; and *less* criticism of their social behaviour; but older students in the UK1 secondary school do seem to follow the teachers' directions much more in comparison to older students represented in the UK27 dataset.

5.41 Comparison of UK1 dataset analysis of teacher-feedback means with UK27 teacher-feedback means

In order to compare the On-task% rates and teacher verbal-feedback means between the UK27 (N=107) and the UK1 (N=129) datasets, an independent samples t-test was conducted. The full original datasets were used to generate means for this procedure:

	UK1orUK27	N	Mean	SD	Std. Err.
On-task%	UK1	129	77.66	14.332	1.262
	UK27	107	78.09	13.260	1.282
INX%	UK1	129	31.88	15.675	1.380
	UK27	107	52.77	18.855	1.823
TPP	UK1	129	6.92	6.272	.552
	UK27	107	16.85	11.083	1.071
SBP	UK1	129	.49	1.724	.152
	UK27	105	2.91	3.732	.364
TPC	UK1	129	.86	1.197	.105
	UK27	107	5.40	6.695	.647
SBC	UK1	129	6.45	6.572	.579
	UK27	107	12.33	9.897	.957

Table 24 – UK27 means compared with UK1 means: 'On-task%'; 'INX%'; 'TPP'; 'SBP'; 'TPC'; 'SBC' using independent samples t-test

A facility of the independent samples t-test using SPSS was that with each comparison of mean-pairs, a choice between an assumption of equal variances and there not being equal variances was legitimised as determined by Levene's Test significance statistic. If $p < .05$, equal variance is *not* assumed, and if $p > .05$ equal variance *is* assumed. The following table displays these considerations and significant results:

	Equal variances	Levene's test for equality of variance		t	df	Sig. (2-tail)	Mean Diff.	Std. Err. Diff.	95% Confidence Interval of Difference	
		F	Sig.						Lower	Upper
On-Task%	Yes	.617	.433	-.234	234	.815	-.424	1.812	-3.994	3.145
INX%	Yes	3.955	.050	-9.297	234	.000	-20.896	2.248	-25.324	-16.468
TPP	No	35.404	.000	-8.237	160.423	.000	-9.928	1.205	-12.308	-7.548
SBP	No	61.264	.000	-6.149	139.826	.000	-2.426	.395	-3.206	-1.646
TPC	No	111.532	.000	6.925	111.633	.000	-4.541	.656	-5.841	-3.242
SBC	No	21.116	.000	-5.256	177.989	.000	-5.877	1.118	-8.084	-3.671

Table 25 – UK27 means compared with UK1 means: 'On-task%'; 'INX%'; 'TPP'; 'SBP'; 'TPC'; 'SBC' using independent samples t-test for equality of means

As the 2 tables show, there were significant differences between UK27 and UK1 'INX%'; 'TPP'; 'SBP'; 'TPC'; and 'SBC' with the means of the UK1 secondary school significantly lower for each variable.

There was no significant difference between UK1 and UK27 'On-task%' rates however. This suggests that in the UK1 secondary school, teachers' verbal-feedback to students as an aggregate: positive comments, criticism; social-behavioural or directed towards academic work-effort; were not considered by teachers to be as strongly associated with whether or not students followed their directions ('On-task%') as they were generally considered to be by UK teachers in the UK27 dataset (see Chapter 6 – Discussion).

5.42 Comparison of UK1 contextual variables with UK27 contextual means

In order to compare the contextual variable means from the UK27 (N=107) and the UK1 (N=129) datasets, an independent samples t-test was conducted. The following variables were considered for inclusion in this analysis:

	UK1 or UK27	N	Mean	SD	Std. Err.
Locality type	UK1	129	2.00	.000	.000
	UK27	107	1.43	.497	.048
Free school meals%	UK1	129	22.26	.000	.000
	UK27	105	20.74	16.579	1.618
Roll	UK1	129	853.00	.000	.000
	UK27	107	959.07	347.893	33.632
Year group	UK1	128	9.53	1.748	.154
	UK27	107	8.66	1.492	.144
Teacher gender	UK1	129	1.60	.492	.043
	UK27	107	1.52	.502	.049
Years teaching	UK1	129	10.37	8.865	.781
	UK27	107	8.18	5.844	.565
Subject	UK1	129	5.03	3.290	.290
	UK27	107	4.30	3.207	.310
Adult present	UK1	129	1.52	.697	.061
	UK27	107	1.34	.752	.073
Class size	UK1	129	17.20	7.710	.679
	UK27	107	19.38	7.071	.684
AM or PM	UK1	129	1.47	.501	.044
	UK27	107	1.32	.468	.045

Table 26 – contextual variable means of UK27 and UK1 datasets

Categorical variables: ’Locality type’; ‘Roll’; ‘Subject’ ‘Year group’ and ‘AM or PM’, were excluded from the independent samples t-test as they were observer-selected in the UK27 dataset. They were of little interest as statistical comparators between the UK27 and UK1 datasets in respect of the research questions of the current research beyond demonstrating that there was not an irreconcilable difference between the numerical ranges of each pair of excluded UK27 and UK1 categorical variable distributions.

Equal variance was tested once again using Levene’s Test significance statistic. If $p < .05$, equal variance was not assumed, and if $p > .05$ equal variance was assumed. The following table displays these considerations and significant results:

	Equal variances	Levene's Test for Equality of Variances		t	df	Sig. (2-tail)	Mean Diff.	Std. Err. Diff.	95% Confidence Interval of Difference	
		F	Sig.						Lower	Upper
Free sch.meals%	No	111.136	.000	.939	104.000	.350	1.518	1.618	-1.690	4.727
Teacher gender	Yes	3.498	.063	1.132	234	.259	.074	.065	-.054	.201
Years teaching	No	8.582	.004	2.278	223.248	.024	2.195	.964	.296	4.093
Adult present	Yes	1.840	.176	1.937	234	.054	.183	.094	-.003	.369
Class size	Yes	.656	.419	-2.246	234	.026	-2.182	.971	-4.095	-.268

Table 27 – UK27 contextual variable means compared with UK1 means: 'Free school means%'; 'Teacher gender'; 'Years teaching'; 'Adults present'; 'Class size', using independent samples t-test for equality of means

5.43 There was no significant difference between the mean 'Free school meals%' of the UK27 dataset of 27.4%; and the figure of 26.2% for the UK1 school.

There was no significant difference between the means of 'Teacher gender' of the UK27 and UK1 datasets. Within the UK27 dataset, 52% of observed teachers were female and 48% male; within the UK1 dataset, 60% of observed teachers were female and 40% male. There was no significant difference between the mean 'Adults present [in the classroom]' during the observed lesson of the UK27 dataset of 1.34; and the figure of 1.53% for the UK1 school.

5.44 There was a significant difference ($p < .05$) between the means of 'Years [teaching] experience' of the UK27 and UK1 datasets. The mean number of years teaching experience of observed-teachers in the UK1 school was 10.37 compared with 8.18 years of experience of observed-teachers in the UK27 dataset. This difference needs to be considered alongside consideration of the significant predictive association made in the analysis of both UK27 ($p < .001$; Chapter 4) and UK1 ($p < .05$; above) datasets between 'Years' experience' and 'On-task%' – meaning students appear to be more likely to be following a teachers' directions, when the teacher has accrued more years of teaching experience.

5.45 There was a significant difference between the mean 'Class size' of the observed lessons of the UK27 dataset: 19.32 students; and the figure of 17.2 students in the observed lessons of the UK1 dataset. This is an unreliable comparison insofar as classes were observer-selected for the UK27 dataset but systematically selected for the UK1 dataset (so

that all teachers in the school were observed twice; once in the morning and once in the afternoon). This procedural difference might explain the UK27 dataset's lower 'Year group' mean (UK27: 8.66; UK1: 9.53) *and* the higher 'Class size' mean. The UK1 procedure meant that a wider range of year-group classes was observed, and therefore, that more small classes in the upper school were observed, as a matter of course. This difference needs to be considered alongside consideration of the significant correlation made in the analysis of the UK1 dataset (*rho*= -.333, N=122, p< .001; above) - but not the UK27 dataset - between 'Class size' and 'On-task%' – meaning students appeared to be more likely to be following teachers' directions when the class-size was smaller in the UK1 school.

5.46 Summative account of an Iterative Learning Conversation with UK1 school using a 3x pass Foucauldian thematic discourse analysis

The discursive analysis of an Iterative Learning Conversation (Apter, 2014) that was held under the auspices of a UK1 school Senior Leadership Team meeting, (ILCS transcription; Appendix) has been summarised here. Thematic considerations[41] from the meeting were recorded, transcribed and annotated using a 3x pass Foucauldian thematic discourse-analysis coding process: pass #1, Iterative Learning Conversations (ILCS) coding; pass #2, Speech Act coding; pass #3, Thematic coding (Chapter 3).

5.47 The UK1 school meeting that was used for this exercise was attended by 8 colleagues: 7 school staff from the school's senior leadership team (SLT) and the facilitator [author]. The meeting lasted 90 minutes.

A transcription of an audio recording of the proceedings of the meeting was made by the facilitator with line numbers separating changes in speaker and subject. Transcribed lines included approximately-timed pauses, repetitions, incomplete clauses, dysfluencies and simple explanatory annotations. The three different coding processes were used by the facilitator to analyse the text *post hoc*.

The analytical standpoint was intentionally the facilitator's. This statement is in keeping with the notion of Foucauldian veracity (Foucault, 1969) where the standpoint of a writer or speaker is made explicit and thus comprehended and taken in to account in the reader's

[41] These summative notes are best read with the colour-coded ILCS transcript of the meeting available to the reader (see Appendix). Line numbers e.g. (line 179): *'I was going to say that….'* refer to the ILCS transcript.

experience of a text. The analytical process was intended to be transparent and the 3 x pass codified transcript is appended (Appendix).

5.48 The meeting was introduced by the Head Teacher but led by the facilitator. The meeting began with the presentation of the initial research findings (uncorrected and prior to examiner-scrutiny). The textual frames of the presentation are appended with the transcript (Appendix). The discourse and the wielding of verbal force within the discourse of the meeting were intentionally provoked by the facilitator's presentation. This was not therefore a free-ranging discussion shaped democratically by all of its members but focussed by the facilitator upon the method, results and utility of the current study from the viewpoint of the school's SLT.

5.49 The major contributors to the discussion, with line by line percentages, excluding the facilitator were: Head Teacher (HT): 38%, Assistant Headteacher 1 (AH1): 7%; Assistant Headteacher 2 (AH2): 14%; Assistant Headteacher 3 (AH3): 14%; Class Teacher 1 (CT1): 5%. Totalling all contributions including the facilitator's presentation together, the facilitator accounted for 56% of contributions and the combined SLT for 44%.

Explicit verbal contributions were made to the meeting by all attendees excepting two female staff, the Head of Key Stage 3 (HKS3) who spoke only to introduce herself at the start of the meeting, and the Head of Key Stage 4 (HKS4) who murmured agreement with other speakers on a number of occasions, and spoke to say (line 179): *'I was going to say that…*' in agreement with AH2's preceding statement. The other female contributor to the meeting was the Head Teacher who made a number of significant contributions and responses within the discussion. Following the Foucauldian rubric, it would be tempting to make a comment from the feminist standpoint about the position of women (HKS3 and HKS4) in the SLT – given also the preponderance of male teachers teaching the upper years in UK1 school, but as this was not a research issue, this would be unnecessarily speculative.

5.50 The initial ILCS and Speech Act coding and percentage-time speaking indicated the overarching positional status of the facilitator as presenter and researcher– insofar as it was the facilitator who determined the agenda, initially asked questions and provoked the direction of the discussion within the meeting. Ten themes were identified in the discourse of the meeting by the facilitator that pertained to the research questions and proposed hypotheses:

No.	*Theme*	Illustrative implicit question
1.	Method	Does this quantitative method with a large sample but small number of variables have research-validity for UK1 SLT?
2.	Historical context and previous research	What is UK1 SLT awareness of previous supportive and competing research in area of classroom observation?
3.	Dependent variable (On-Task)	Is student on-task behaviour perceived by UK1 SLT to be a robust-enough dependent variable to focus a classroom observation upon?
4.	Independent: teacher verbal behaviour	Is teacher verbal feedback in 4 categories: academic positive, academic criticism, behaviour positive, behaviour criticism, perceived by UK1 SLT to be a robust-enough set of independent variables?
5.	Independent variables: non-verbal	Do any emergent correlations or causal associations with on-task behaviour have face-validity as perceived by the UK1 SLT as perceived by the UK1 SLT?
5.	UK1 dataset: results and solutions	Does the UK1 dataset accurately represent the social-emotional climate for learning in academic classes in the UK1 school as perceived by the UK1 SLT?
7.	UK27 dataset: results	Does the UK27 dataset have face-validity as representing the social-emotional climate for learning in academic classes in UK secondary schools generally?
8.	UK1 v. UK27 datasets: comparison	Does a comparison of the 2 datasets have validity and utility for the UK1 school SLT?
9.	Ofsted	What effects have the Ofsted inspection-judgements had on the UK1 school and what utility for the UK1 SLT might the current research results have in responding to Ofsted challenges?
10.	Omissions and alternative (non-research) solutions	What has been missed out in the current research and what solutions might be available to the school that are not indicated by the current research as perceived by the UK1 SLT?

Table 28 - Ten themes used for Foucauldian 3-pass thematic discourse-analysis (Apter, 2014; Harding 2013)

It is also evident that from this analysis that the facilitator's discursive effort was to convince the SLT of the research's standpoint-position – that the current research, the derived datasets and the facilitator's analysis was significant, and had significant utility for the SLT and their school. This was not a neutrally delivered presentation therefore. The responses of members of the SLT to this provocation in the meeting provided the information within the transcription that was subsequently thematically coded.

5.51 A tipping-point (line 79) in the discourse occurs revealing a move away from a unanimous compliance of the SLT with the facilitator's standpoint as it has been

represented in the opening presentation. This occurred when a male Assistant Head Teacher, AH2 made an unexpected 'grab for power' in Foucauldian terms (Foucault, 1969) using a series of clarifying questions, beginning with the deferential: '*You are probably going to get to this… I am probably jumping the gun…'*

Later (line 179), AH2 passively voices again his doubt that the psychologist-researcher's tight focus on a small number of variables was unlikely to produce the rich data that a more wide-angled view would produce. The scepticism that AH2 voices is that a qualitative multivariate approach to these observations would have produced a richer picture than the narrow quantitative approach that the facilitator had used. This view is developed by a Class Teacher SLT representative, CT1, who reflects (line 180) that a qualitative approach to data-collection could have been used. This criticism is responded to in Chapter 2 – Methodology and Method.

5.52 A tipping-point back towards the research's standpoint position and the conclusion of the discursive process is revealed at the end of the meeting by the Head Teacher, HT: *'But I definitely want to address the way… And how much… We talk to students… We do need to get better at this… And we'd like to have you back… Maybe to help Sean… Re-run this* [research] *with the maths department… To show how much they have moved on…'* (line 184)

5.53 The discursive process between the two tipping points in Foucauldian terms in respect of the use of verbal forcefulness appeared to move from compliance with the facilitator's presentation and hypothesis to challenge (line 79) through an exercise of scepticism in respect of method (lines 79; 132; 189).

This modified compliance might be expressed thus: deficiencies in the use of teacher verbal feedback need to be addressed in UK1 school as a matter of urgency but with the caveat that it was not as simple as increasing or decreasing different types of verbal feedback. Points in developing this modified compliance are revealed (line 104): 'AH2: *It makes sense…*'; (line 129): *'It's interesting when we talk about… Er… Verbal feedback… And generally the feedback as a whole across the school… Is that whether we are challenging enough or supportive and pushing forward enough… And higher attainers… Just seeing students are on-task or not and whether we are happy* [complacent] *with that or not… I think praise is a very complex thing, isn't it…'*; (line 147): '*Some of them who are good mathematicians don't know how to communicate with adults… Never mind children...'*; (line 156): '*There's*

a lot of work to be done with just the simple… Adults… And how they converse with children… In school…'; (line 157): *'Rather than when we are talking verbally… The teachers are talking verbally… How is it… How is it… Er… Measured as the student verbal feedback? In pushing their progress forward… In the verbal context…'*

5.54 The discursive process and the stages of response-transformation that resulted in the second tipping point and the move of the SLT back to a modified compliance with the standpoint position of the current research (line 184) appeared to be motivated by the impact of the maths / English dichotomy as it was revealed by the facilitator's analysis of the UK1 dataset.

Indeed, the UK1 school SLT appeared less impressed and affected by the UK1 UK27 dataset dichotomy than this Maths / English dichotomy (introduced: lines 101 – 103); a phenomenon that the school were sensitively aware of after the similar judgement of Ofsted in 2014 (lines 104; 106; 144 – 147; 184). The single factor that had ensured that Ofsted could make a 'failing school' judgement had been the poor performance of the maths department according to the Head Teacher. This means that there was an apparent broad agreement about this between the facilitator's presentation of the analysis of the UK1 dataset, the perceptions of the UK1 SLT and the observations of Ofsted. However, in the detail of those three viewpoints and the respective processes that those viewpoints had been arrived at, there were many significant differences (lines 101 to 104; 112; 144; 145; 147).

5.55 Whilst the Head Teacher of UK1 school appeared to be enthusiastic about the current research, the UK1 dataset and the implications for management decisions, it was not so evident that the SLT were as in favour of a psychologist's view of the school's predicament (lines 149; 173; 179) but rather preferred the advice of educationalist-advisers who had previously been Head Teachers.

5.56 Sometimes members of the SLT appeared critical of the Head Teacher's leadership (lines 136; and 151). The Head Teacher was not directly responsive to these criticisms but says (line 142): *'From a leadership perspective, from our knowledge of what is going on. It's a lot tighter. The data and the understanding of how children are progressing... Is much tighter… Since John and Chris* [first names of Assistant Head Teachers, AH1 and AH2 – the HT's critics] *have been looking at that.'* Later, the HT brings together the achievements that she is claiming, the efforts of AH1 and AH2, and the efforts of the psychologist-researcher (line 144): *'But it's useful; though because our English data on its own entirely*

confirms your data… Because last year… Those children… Who would have been part of that… Achieved the best set of results we've had…'

The facilitator and the SLT were in agreement that Maths teachers in the school were not effective in using feedback (lines 104; 106) and were thus unable to communicate an enjoyment of the subject and a love of learning within that subject, whereas Ofsted described the failure simply in terms of measured outcomes. This appears to lead in to the consensus that is voiced at the end of the meeting by the Head Teacher (line 184).

5.57 Shortly after the observations for the current study had been completed in UK1 school, and after a number of Ofsted follow-up visits had been completed that judged the school not to be making sufficient progress, the school and their governing body submitted to Ofsted's pressure (lines 34-36; 138; 169; 172) to become an 'academy' (see also Mansell, 2011).

This development meant that the 'special measures' imposed by Ofsted and the subsidiary follow-up inspection visits ceased. Under current legislation (DfE, 2015) a school that elects to become an academy becomes a new school with a new name. Their failed status and the special measures that have been imposed are revoked. Their Ofsted inspection cycle begins again from the date that the school becomes an academy.

It was understood that the UK1 school was told that they would not have their first Ofsted inspection under the auspices of their new name and academy status for at least 18 months measured from the date of the observations used in the current research (line 142). This must be viewed as a strong incentive for any school to become an academy and to comply with all guidance given them by Ofsted given the punitively-perceived affect of being judged to require 'special measures' (lines 138 to 140).

5.58 The conclusion of the discussion within the meeting was that the current research, the UK1 dataset and the comparison with the UK27 dataset was a worthy exercise with utility for the SLT and their school.

However, the impending re-visit of an Ofsted inspection team and observation of lessons would be primarily concerned with a range of factors that were not included in the current research, e.g. learning outcomes and evidence of curricular knowledge-learning within lessons.

The SLT's shared view and particularly the view of the Head Teacher was that the way that teachers in the school gave verbal-feedback to students in all year groups and curricular areas but most particularly in maths needed to be measured, monitored and significantly improved.

The issue of how teachers used verbal feedback in the UK1 school had been recognised as an issue by the Head Teacher and this had motivated her to commission this part of the current research: the UK1 dataset and its analysis.

The Ofsted inspection and judgement had superseded the UK1 school's best effort to address this matter independently through commissioning the research. Unfortunately, improved teacher-feedback to students was only one of many managerial measures that the SLT needed to embed in the school prior to the Ofsted revisit which was due to happen in the next few months (lines 157; 168; 172-173; 184).

Chapter 6 – Discussion around research questions, hypotheses and findings: critiques, responses and suggested revisions

- **Research Question #1 – Hypotheses #1; #2; and #3**
- **Research Question #2 – Hypotheses #4; #5; #6; #7; #8; #9; #10; #11; #12; #13**
- **Research Question #3**
- **Research Question #4**
- **Research Question #5**
- **Research Question #6**

6.01 Research Question #1

Research Question #1 was: 'What types of Teachers' verbal feedback encourage students to follow their directions and stay 'on-task' in academic lessons in UK secondary schools?' The estimation of the extent to which the results of this research addressed this question leads to complex and tentative conclusions.

It was evident from the results (Chapters 4 and 5) obtained in this research and from the statistical analysis of the UK27 and UK1 datasets provided here that the *causality* of the association between students' following their teachers' directions ('on-task' behaviour) and any specified type of verbal feedback used by teachers in academic classrooms in UK secondary schools can only be speculated about. Whilst with both datasets a significant association was found between teachers' use of social behavioural criticism ('SBC') and students' following teachers' directions ('On-task%'), which was 'stimulus' and which was 'response', or whether the direction of the causal relationship changes for different exchanges, was an arbitrary judgement because of the temporal organisation of the time-sampling method used by the *MICRO* observational schedule.

6.02 Received statistical wisdom (Field, 2009; Palint, 2010; Tabachnick and Fidell, 2012) suggest that multiple regression can be used to establish causal relationships between one or more independent variables and a dependent variable. It is contended here that this cannot be done whilst ignoring a structural and temporal disconnect that exists between the actual behaviour; and the recording of the observation of an indicative sample of similar behaviours as has been the procedure with time-sampling methods used in the current research; and with similar mixed sampling methods employed by previous structured-observation feedback-researchers that have been cited here (White, 1975; Persons, *et al.*, 1976; Nafpaktitus, *et al.*, 1985; Wheldall, *et al.*, 1989; Harrop and Swinson, 2000; Swinson and Harrop, 2001; Apter, *et al.* 2010).

The firmest conclusion that can be made from the findings reported here is about the probability of bi-directionality of influence in the associations that exist between student's following teachers' directions and teachers' verbal feedback.

It is meaningful nonetheless to consider the nature of these associations in providing a simple quantified indication of the more generally specified social-emotional climate for learning in classrooms – as suggested by the author, and how future versions of *MICRO* might be better designed to describe the conduciveness of this climate.

6.03 Hypothesis #1

Research question #1 provided the auspice for the first 3 hypotheses that were examined.

Hypothesis #1 proposed that: 'Higher frequencies of *positive* verbal feedback by teachers in academic lessons in UK secondary schools directed towards their productive academic-task-focussed behaviour and product ('TPP'), e.g. written work, is *causally* associated with the time spent by students following a teacher's directions - as measured using a quantified dependent variable ('On-task%).'

No significant evidence was found to support Hypothesis #1 in academic lessons in UK secondary schools (Chapter 4; Chapter 5) using HMRA. It was also found that 'TPP' was *not* correlated with 'On-task%' using Spearman's rank ordered correlation *rho* in either the UK27 and UK1 datasets.

This finding contrasts with the author's findings in primary schools. In the primary study, positive recognition of task-performance was significantly correlated with on-task behaviour (Apter, *et al.*, 2010, p.155). It was also contrary to the results of previous research in secondary schools (e.g. Wheldall, *et al..*, 1989; Harrop and Swinson, 2000) and in primary schools (Merrett and Wheldall, 1987; Harrop and Swinson, 2000).

It is probable that the 'paradoxical effect' of praise-type feedback and its propensity to limit the achievement of some students (Dweck, 1986; Meyer, 1992) would be most pronounced in the secondary school because of the age of students. It is easy to accept Dweck's (1999) advice about the most effective praise statements as the current research was not concerned with the linguistic detail of 'praise' statements in the UK, but anecdotally, UK teachers in

both primary and secondary schools do not praise students – excepting the very young – by telling them that they are 'clever' or 'smart'. This might be a North American cultural artefact, telling older students that they are 'smart', and less relevant to the UK secondary context.

Indeed, it might be the case that UK secondary teachers believe that it is best to avoid all public praising of students, particularly in the upper years. This would require further investigation of secondary teachers' beliefs and attitudes. In contrast, primary students appear to usually enjoy and are motivated by the public celebration of their academic successes by their teachers (Apter, *et al.,* 2010) and primary teachers evidently act on this.

The analysis of both UK27 and UK1 datasets shows that teachers use more feedback in Year 7 than they do in successive years through to Years 11 and 12, as if they might be adjusting their use of feedback according to the age of the students that they are teaching. Speculation about how age affects the way in which approval and other types of feedback are perceived by teachers to be received by students was first written about by White (1975) in her seminal teacher-feedback study. Further investigation could involve further research into the phenomenological experiences of teachers and students in both school phases.

However, it is still controvertible as to why the result obtained from this current large-scale study of secondary schools is so different from findings from previous secondary school research (e.g. Wheldall and Merrett, 1989; Harrop and Swinson, 2000) in finding that teachers in secondary schools generally do not use significant amounts of any type of non-private (observable) verbal-feedback comments in their academic classrooms other than when they criticise unwanted social-behaviour.

Further consideration as to why this hypothesis (#1) was not supported by the findings of the current research is provided under the discussion (below) of Hypothesis #3.

6.04 Hypothesis #2

The second hypothesis under the auspice of the first research question, Hypothesis #2 proposed that: 'High levels of time following a teacher's directions in secondary classrooms ('On-task%') are associated with teachers who use more verbal teaching behaviour' ('INX%'), i.e. students are more on-task when teachers talk more in lessons, was not supported by analysis of either the UK27 or the UK1 datasets.'

This hypothesis was adopted *a priori* for the current research from the author's primary research (Apter, *et al.*, 2010). It was an unexpected finding when the primary school research was completed but it became the finding that attracted the most press interest at that time[42].

With the primary school study (ibid.), 'ARD' (academic or routine direction) and 'On-task' were weakly but significantly correlated (r= .168, N=137, p< 0.05, 2-tailed) but in the current study using Spearman's rank ordered correlation coefficient (*rho*) 'INX%' and 'On-task%' were not significantly correlated in either the UK27 or the UK1 dataset.

Unexpectedly, the current research has shown that teachers in academic lessons in secondary schools speak more on average ('INX%'= 52.66%) when teaching academic lessons than did their primary school counterparts ('ARD'= 22.5%, ibid.) but as the observational schedules in respect of 'ARD' and 'INX%' differ slightly, it cannot be argued that the measures are equally valid or that a comparison can be made with adequate precision. Indeed, all measures of feedback ('TAP'; 'TAN'; 'TSP' and 'TSN') and neutral teacher-talk ('ARD') were standardised to 'rates per minute' in the published report of the primary study (ibid., p.155). In the current study, raw scores were not standardised (see below).

One of the reasons for the high teacher-talk rate (INX%) found in the current research might be that secondary curricular material might require more verbal explanation dependent on subject, student age and / or on the teaching approach, but initial investigation of the 'Subject' and 'Year group' variables suggested that this variation had little or no bearing on student engagement in the lesson. This new hypothesis would require a new investigation.

6.05 Hypothesis #3

The third hypothesis under the auspice of research question #1, Hypothesis #3, proposing that: **'**teachers in secondary schools use significantly more *positive* verbal feedback, e.g. praising comments, to *encourage* academic-task-focussed behaviour and product; and significantly more *negative* verbal feedback, e.g. critical comments, to discourage undesired

[42] Pupils 'behave better than 1970s' by Sean Coughlan – BBC Online http://news.bbc.co.uk/1/hi/education/7745772.stm

social-behaviour ('Off-task' behaviour); and use significantly less *positive* verbal feedback to encourage students to follow their social-behavioural (non-academic) directions; and significantly less *negative* verbal feedback to discourage undesired, incorrect or inaccurate academic-task-focussed behaviour or product', is clearly supported by the findings.

Both datasets, UK27 and UK1, support the hypothesis (#3) showing ratios going in the anticipated direction.

The following table presents data aggregated from the 2 datasets and previously presented separately in more detail (Chapters 4 and 5):

Dataset	TPP	Std. Dev.	SBP	Std. Dev.	TPC	Std. Dev.	SBC	Std. Dev.
UK27 (N= 106)	16.78	11.083	2.89	3.72	5.11	6.02	12.21	9.87
UK1 (N= 122)	6.63	6.088	.19	.433	.86	1.208	6.42	6.558

Table 29 – comparison of teacher feedback types - means and standard deviations

The data shows that 'TPP' in both datasets is significantly larger than 'SBP', 'TPC' and 'SBC'; 'SBC' in both datasets is significantly larger than 'SBP' and 'TPC'; and 'TPC' in both datasets is significantly[43] larger than 'SBP'.

However, the face validity of these differences, and the ratios obtained (Chapters 4 and 5) does not take account of the number of zero-scores in each category – observations of lessons where teachers did not use any feedback comments of one category or another:

Dataset	TPP zeros	SBP zeros	TPC zeros	SBC zeros
UK27 (N= 106)	3	38	14	6
UK1 (N= 122)	11	101	73	17

Table 30 – comparison of total number of data points and zero scores in the UK27 and UK1 datasets

[43] Data from ANOVA showing significant differences is presented more fully in Chapters 4 and 5.

This table emphasises the finding that teachers in the UK1 school used all 4 types of verbal feedback behaviour significantly less than the UK27 dataset.

This table also amplifies the finding discussed previously that 'SBP' and 'TPC' feedback types – respectively: approval of social behaviour; and criticism of academic-task behaviour; were often unused by teachers in secondary school academic lessons according to observers.

6.06 A proven *post hoc* hypothesis proposed that the direction of the significantly predictive relationship between 'SBC' and 'On-task%' was as follows: that the less a teacher was observed to use social behaviour criticism ('SBC'), the more 'On-task' the students in their class would be expected to be. There would appear to be an obvious face-validity to the causal-direction of this association (Chapter 4).

There is also a more novel way of conceiving of the relationship between the 2 variables statistically: that a teacher's social behaviour criticism ('SBC') is the response to student 'Off-task' behaviour, and could thus be treated in a separate analyses as the dependent variable with 'Off-task' being treated as an independent behaviour.

The bi-directionality of influence in the associations between dependent ('On-task%') and independent variables (e.g. 'SBC') suggests that maybe the choice of multiple regression was unnecessary when correlational statistics and scatter plots would have provided as much usable statistical power as was necessary - as was concluded with the author's primary-phase study (Apter, *et al.* 2010) by the peer review process.

6.07 Particularly, the observations of teachers in the UK1 school as they are recorded in this table show that in the majority of academic lessons, teachers did not use positive recognition of social behaviour ('SBP'; telling students that they were behaving well when they were) and public criticism of students' task-performance criticism ('TPC'; telling students that they are not working on their academic work-tasks in an effortful or correct way). As strong predictive associations have not been summatively made between either of these feedback categories and 'On-task' behaviour; and in consideration of the finding that there was not a significant difference between the 'On-task%' rates of the UK27 schools' and UK1 school's datasets, any response requiring teachers in the UK1 school to use more of these 2 types of feedback is pre-empted pending further investigation.

Given the low frequencies of teacher feedback that were *not* criticisms of behaviour ('SBC') in both UK1 and UK27 datasets, it would appear that teachers working contemporarily in secondary schools teaching academic lessons rarely used any of the other 3 types of feedback, categorised as such for this study ('TPP'; 'TPC'; and 'SBP'; though differentially, slightly more 'TPP', compared to the other two).

6.08 The apparently higher usage of these 3 types of feedback in previous secondary school studies (Wheldall, *et al..,* 1989; Harrop and Swinson, 2000) might be considered a historical artefact, given that considerable time has elapsed (10 years or more between Harrop and Swinson, 2000, and when data was collected for the current study) but this is debatable without a robust causal explanation as to how classroom discourse has changed in that time. This would require further investigative research to establish.

From their own study of 10 infant teachers, 10 junior teachers and 10 secondary teachers, Harrop and Swinson (2000) reported that there were no significant differences in the use of approval by teachers between infant, junior and secondary schools. In contrast, they reported that there was a decreasing use of disapproving comments by teachers from infant through to upper-secondary. It is difficult to see why the findings of the current research reported here are so different to the secondary school results reported by Harrop and Swinson (ibid.) but a number of factors are worth considering.

For their study, Harrop and Swinson (ibid.) standardised scores from raw tallies and arrived at 'rates per minute' when considering the use of the 4 different types of feedback. Wheldall, *et al.* (1989) also standardised raw tallies of comments into frequency rates in their secondary study. As noted above, this statistical method was used with the author's primary study (Apter, *et al.,* 2010). Standardisation in this way means that the real-time ranges of raw scores cannot be extrapolated by the reader. As a standardisation process was not used on the raw scores of the current research, it became obvious that the differences in frequency ranges of the 4 different types of feedback ('TPP'; 'TPC'; 'SBP; 'SBC') were so significant that if these were ignored, injudicious hypothesising would result in a Type 1 error.

It is unlikely that the low frequencies' of the 4 feedback types reported here as part of the current research in secondary schools can be judged to be an unwanted observer-effect because similarly low and narrow ranges occurred within both datasets: the UK27 dataset

with 29 different observers; and UK1 dataset with 4 observers (supervised by, and including, the author).

6.09 It is unlikely too that the low frequencies (of 'TPP'; 'TPC'; 'SBP') and high numbers of zero scores were caused by any other type of methodological confoundment such as inadequate training and preparation of observers, as the expectation would be for inadequately prepared observers to over-estimate frequencies of this type of data, rather than under-estimate (Rapp, Carroll, Swanson and Higgins, 2015). For example, observer expectancies would predict up-rating of marginal instances of teachers' verbal feedback comments at least as much into the 'TPP' category as into the 'SBC' category.

An anticipated Hawthorne-type observer-effect (Draper, 2014) that would suggest that the reported frequencies are higher than would otherwise be the case, particularly in respect of 'TPP' and 'SBP', is that teachers might have been wanting to impress psychologist-observers by being more positive than they usually were with their students, but the low and narrow range positive comments tallies ('TPP' and 'SBP') indicated that this was not the case. Indeed, the primary data that the author collected (Apter, *et al.*, 2010) where the raw unstandardised frequencies of 'TPP' (termed 'TAP' in the primary research report; ibid.) and 'SBC' (termed 'TSN' in the primary research report; ibid.) were similar to those revealed in the current study. This provides further support to the argument that the *MICRO* method and any observer-effects that were entailed by following the method were not responsible for the low frequencies of non-'SBC' feedback variables however.

6.10 It is not clear as to whether the finding that teachers used little feedback in academic lessons in secondary schools might have also been the case with previous secondary school research (e.g. Wheldall and Merrett, 1989; Harrop and Swinson, 2000) as raw data from those studies is not available. If a behavioural phenomenon is not a real or realisable possibility as a significantly frequent event in a given context because it is indistinctive to an observer; or because it is inadequately operationally-defined; or because it is not ever used; or cannot be functionally used in its specified form and context, the claim of significance of ratio-difference with another much more frequent behavioural phenomenon where any (or all) opposite conditions apply: it is observably distinctive; adequately operationally-defined; can be functionally used; and is frequently observed to be used, is implausible.

To illustrate with an example: if it were planned that a UK teacher was to be observed using two types of feedback: using the positive verbal comment in English, 'Well done!' and using the Inuit words for 'not good enough', it is probable that a resultant reported ratio would appear to suggest that the positive verbal comment 'Well done!' was used much more frequently than the critical phrase 'not good enough' in Inuit: 'Piktaungitok!' However, the real obstacle to the UK teacher using the critical feedback 'not good enough' in Inuit is that the UK teacher does not know any Inuit, and even if they did, it would not be contextually comprehensible or culturally appropriate to use such feedback in a UK secondary classroom as it would not be understood by the students.

6.11 There were a number of correlations that were calculated using a non-parametric test[44], Spearman's rank order coefficient *rho*, in respect of teachers' verbal feedback to students that were revealed during the statistical analysis of the UK27 and UK1 datasets that were neither anticipated nor hypothesised, but which have been reported as emergent findings (Chapter 4 and 5) for the sake of completeness. To recap: the following correlations were produced from both datasets: 'TPP' and 'SBP' (UK27: $p < .01$; UK1: $p < .05$); 'TPP' and 'TPC' (UK27: $p < .05$; UK1: $p < .05$); 'TPP' and 'SBC' (UK27: $p < .05$; UK1: $p < .01$); and 'SBC' and 'TPC' (UK27: $p < .01$; UK1: $p < .05$).

It could be concluded from these correlations that teachers who use more positive verbal feedback directed towards a student's academic work ('TPP') are most likely to also positively recognise 'good' behaviour ('SBP'); criticise academic work ('TPC'), and criticise 'bad' behaviour ('SBC'); and teachers who use more critical verbal feedback directed towards students' (social) behaviour are most likely to both publicly approve of, and criticise a student's academic work ('TPP' and 'TPC'). It is important to remember the caveats that raw scores for 'TPP', and more so for 'TPC' and 'SBP', were low in number and narrow in range in the UK27 dataset and lower still in the UK1 dataset; and the number of non-zero tallies diminished and the range narrowed further in the upper years in both datasets (Chapter 4 and Chapter 5).

6.12 Within the field of teacher feedback research, Nafpaktitus, *et al.* (1985) emphasised the dangers of making too much of correlations in classroom observations of teacher feedback:

[44] There was also a significant parametric correlation in the UK27 dataset using Pearson's product moment *r* between 'TPP' and 'INX%' ($p < .05$); and a non-parametric correlation in the UK1 dataset using Spearman's rank order *rho* between 'SBP' and 'INX%' ($p < .05$) but as neither of these findings were replicated between the two datasets, they have not been expanded upon here.

"Caution should be used in interpreting correlations, given that the more correlations generated, the greater the probability of obtaining significance by chance. (p.367)

Recently, there has been considerable concern voiced in the psychology literature and in the wider general science literature about the frequency of Type 1 errors and doubts about the replicability of psychological experiments (Baguley, 2010; Pashler and Wagenmakers, 2012; Open Science Collaboration, 2015). It would seem wise to proceed with caution when discussing correlations revealed in the current research and drawing conclusions from them.

6.13 Research question #2

It was intended that a range of assumedly influential contextual variables in secondary schools would be addressed and examined by Research question #2: **'**What other factors are significantly associated with students following teachers' directions ('On-task' behaviour) in academic lessons in UK secondary schools?' Significant findings about each contextual variable are discussed in-turn below.

6.14 Each hypothesis under the auspice of research question #2: Hypothesis #4, #5, #6, #7, #8, #9, #10, #11, #12 and #13, was of the same form: the degree to which students follow a teacher's directions in secondary classrooms ('On-task%') is significantly associated with: 'Locality type' (hypothesis #4); 'Free school meals %' (#5); 'School roll' (#6); 'Year group' (#7); 'Teacher gender' (#8); 'Years [of teaching] experience' (#9); 'Subject' (#10); 'Adults present' (#11); 'Class size' (#12)'; 'AM or PM' (#13). The direction of influence was not hypothesised. The null hypothesis in each case would be that there was no significant association between the dependent variable and the contextual variable being considered.

Statistically significant results were found from using hierarchical multiple regression analysis (HMRA) in respect of predictive associations between the distributions of the dependent variable: students following teachers' directions ('On-task%), and: 'Free school meals %' in the UK27 dataset; 'Years' experience' in both the UK27 and UK1 datasets; and 'Class size' and 'Year group' in the UK1 dataset.

There were no significant associations revealed using HMRA between: 'Locality type'; 'School roll'; 'Teacher gender'; 'Adults present'; 'AM or PM'; and the dependent variable 'On-task%' in either the UK27 or UK1 datasets[45].

6.15 'Free school meals%' was found to be individually significant in predicting the distribution of the 'On-task%' variable within the UK27 dataset ($p < .05$; Chapter 4). It seems an unlikely association but this statistic appeared to suggest that the more free school meals were registered in a school, the more students would be found to be following teachers' directions in that school. The UK1 school had 22.26% of students registered for free school meals ('FSM%') from a roll of 980, compared to a national figure (DfE, 2015c) of 14.6% in 2014.

The mechanical linkage between 'Free school meals%' and 'On-task%' is not obvious and there is therefore a possibility of a Type 1 error if too much significance is made of such a predictive association. More investigation of this association would need to be completed before a more meaningful hypothesis could be generated.

The debate about whether or not the percentage of free school meals registered in a school is a valid indicator of poverty has largely concluded with academics and teachers deciding that is an unreliable indicator of poverty – primarily because of regional and cultural variation in take-up of the offer (Hobbs and Vignoles, 2007; Kounali, Robinson, Goldstein and Lauder, 2014).

6.16 The positive and predictive direction of the association between the number of years that a teacher has been teaching ('Years' experience') and the degree to which students follow their directions in class ('On task%') appears to have a more plausible and meaningful connection; and the finding was replicated in both datasets (HMRA; Chapter 4 and 5). The more experienced a teacher, the more likely that students will follow their directions in academic lessons – presumably because either: the teacher is more likely to have learnt how to teach effectively and keep the students engaged; or, only teachers who are effective in the classroom stay in the profession for many years; or a combination of these two factors.

[45] 'Subject' was not included in the HMRA, and is discussed separately.

Considering non-parametric correlations in the UK27 dataset using Spearman's *rho*, the 'Years experience' distribution was also correlated with the amount the teacher spoke in lessons ('INX%'; *rho*= .239, N=106, *p*= <.05) and negatively with the amount of critical feedback they directed at the social behaviour of students ('SBC'; *rho*= -.220, N=106, *p*< .05) - meaning the more experienced the teacher was, the more they would talk but the less they would criticise students' behaviour in academic lessons. Once again, the reasons as to why these associations were found, and what the mechanical linkage was, and in which direction the causal vector was operating, appears opaque and further investigation would be necessary to clarify what was happening.

6.17 Within the UK1 dataset 'Class size' was revealed by HMRA to be predictively associated with 'On-task%' and individually significant (*B*= -.250, *p*< .05) but not within the UK27 dataset. The face validity of concluding that students were more likely to follow teachers' directions in smaller classes appears plausible. The equal spread of collected data throughout the UK1 school, year by year, might explain why the result was significant; in comparison to the UK27 dataset, where there were a disproportionate number of Year 7 and 8 observations (Chapter 4) where class sizes were larger.

However, smaller class size has repeatedly been debunked as a desirable development by Ofsted and the DfE. Ofsted reported in 1995: "The main finding of the report, that reductions in class size do not necessarily lead to better teaching and higher standards, suggests, however that schools should not automatically seek to use new resources to reduce class size." [46] (p. 44)

The study by Hattie (2009), quoted by the DfE (2011), found the 'effect size' of reducing class size on attainment to be smaller than a number of other interventions. Hattie argued that value for money in raising attainment in schools is better achieved through other interventions rather than class-size reduction. Hattie did not deny that the balance of evidence from 3 meta-analyses and 96 studies showing that there were significant benefits to be had from smaller class sizes, but that it was much less cost-effective than 105 other 'interventions'.

[46] In the same report the authors note that to reduce the average class size in primary schools by one student at 1995 prices would cost £170 million.

However, psychological first principles and a large number of evidential meta-analyses show how children learn best with individualised feedback (Hattie, 2009). Hattie also reported that individualised teaching approaches worked to promote students' learning and well-being (ibid.). The benefit of smaller class-sizes that facilitate more individual teaching has been approvingly researched in North America (Whitehurst and Chingos, 2011). The benefits of individual tutoring is also promoted by the Sutton Trust (2015). Obviously, a smaller class facilitates individual tutoring by the teacher during a lesson. This is presumably why elite private schools usually opt for smaller class sizes[47]. All of this might promote some scepticism of the Ofsted and DfE mantra that reducing average class-sizes has little effect upon educational outcomes.

6.18 Within the UK1 dataset, but not the UK27 dataset, 'Year group' was revealed by HMRA to be predictively associated with 'On-task%' and individually significant ($\mathcal{B}$= .238, p< .05). Again, the face validity of concluding that older students in smaller classes were more likely to be following the teacher's directions ('On-task%') appears reasonable; the disproportionate number of Year 7 and 8 observations in the UK27 dataset might explain why this result was not replicated.

6.19 A particular contextual variable ('Subject') became more important for the single secondary school used for the UK1 dataset, and that was a comparison of the performances of teachers and pupils in maths and English lessons in all year groups. This was because the school's maths department had been particularly criticised by Ofsted and this had been a contributing factor when they had placed the school in 'special measures' at the time that the UK1 dataset was being collected. In contrast, the English department had been positively recognised and celebrated[48].

The importance that the senior leadership team (SLT) of the UK1 school attributed to the difference between the results obtained by this study in respect of observations of English

[47] Maximum primary class size in New Zealand, where Professor John Hattie works is set by statute at 25. In the UK, the maximum is set by statute at 30 students. The ratio of teachers to students in the UK private primary sector was 1:15 compared to an average of 1:26 in the state sector; and 1:18 for private secondary schools compared to 1:21 for UK state secondary schools (Paton, 2014; OECD, 2012). Private schools in the UK, e.g. Eton College, where the cost of employing more teachers is less of an obstacle, celebrate smaller class sizes in their publicity: Eton primary department quote maximum class sizes: 20 to 24; and Eton lower-secondary department class size: 12 students.

[48] It emerged from discussions that the author had with the Head Teacher of UK1 school at the time that going in to the Ofsted inspection in December, 2013, the Senior Leadership Team of the school had been anticipating this criticism of maths teaching in the school (see ILCS transcript, Appendix).

and of maths lessons is evident from the ILCS transcript (Appendix) and the analytical summary of the same (Chapter 5).

Students in maths lessons in the UK1 secondary school followed teachers' directions much less than in English lessons. In English, teachers and students managed to establish higher levels of students following teachers' directions in lessons in Year 7; and this also increased, year on year, into the upper school[49].

In the UK27 dataset, no significant differences were noted between teachers' usage of different types of verbal feedback between English, maths and science lessons, and also between the levels of 'On-task%' in different subjects. In the UK1 dataset, 'On-task%' was significantly higher in English lessons compared to maths lessons ($p < .05$; Chapter 5), but there was no significant difference in 'TPP' and 'SBC' between English and maths. The only significant difference was an increase in the observed use of critical verbal feedback directed at academic work ('TPC') in maths.

The differences between findings from observations of English and maths lessons in the UK1 dataset might be more influenced by curricular materials or by differences in departmental leadership - factors unexamined in the current research – rather than by patterns of teacher feedback. However, the ILCS transcript (Appendix) and summary (Chapter 5) revealed that the Head Teacher and SLT believed that it was individual teacher factors that were responsible (ILCS transcript, line 147, Appendix). The UK1 SLT believed that teacher feedback was something that needed to be improved in terms of both quantity and quality in lessons in the school in all years; and in all academic subjects except English (ILCS transcript, lines 157; 168; 172-173; 184).

6.20 There were a number of emergent un-hypothesised findings that should arguably be emphasised in respect of some further associations that were found between teacher's verbal feedback, contextual variables, and students following teachers' directions ('On-task%').

In both UK27 and UK1 secondary school datasets, when teaching older students, e.g. years, 10, 11 and 12, teachers used less positive comments directed towards students' work

[49] Excepting a brief dip in Year 8 – described as a 'difficult year group' by the UK1 Head Teacher and the senior leadership team.

('TPP') and fewer critical comments directed towards their social behaviour; but students were also less likely to be following teacher's directions ('On-task%'). This was intriguing, for while no generally predicative association has been found between 'TPP' and 'On-task%' in the results presented here (Chapters 4 and 5), a significant predictive relationship has been found between 'SBC' and 'On-task%'. This invites the question: why are teachers of years 10, 11 and 12 academic lessons not using 'SBC' more in order to increase the compliance and engagement of students? In order to understand this better, it might require further investigation focussing upon the teaching theories and history of these teachers.

It appeared from correlations between contextual variables from the UK1 dataset that male teachers within that secondary school were more likely than female teachers to be teaching the upper years, have more years teaching experience, and be teaching smaller classes (Chapter 5). These findings were not replicated in the UK27 dataset.

6.21 Research question #3

Research question #3 was: 'Is the *MICRO* observational tool a valid and reliable means for measuring the association between a teacher's verbal behaviour; and the length of periods of time students spend following a teacher's directions in classrooms in UK secondary schools?'

On completing this research, it was always going to be important to re-evaluate the *MICRO* observational tool as a valid and reliable means for measuring the association between a teacher's verbal behaviour; and the length of periods of time students spend following a teacher's directions in classrooms in UK secondary schools.

There were a number of issues with the method chosen that would compel revision of the research design and of the *MICRO* schedule. Redesigning *MICRO* would prove advantageous by broadening its scope and utility; by improving its validity; and by further ensuring its reliability.

6.22 A number of observer-effects have been discussed above. A revision of the *MICRO* method would have to consider and address these effects. As has been discussed above, all raw scores and percentages collected for both the UK1 dataset and the UK27 dataset were probably higher to a small extent than unobserved unobservable theoretical baselines for each lesson, but as the research in its current methodological form mainly depended upon

comparisons of different measures (between-measures) and different groups (between groups), findings have the validity-strength of comparing like with like. This might be anticipated to cancel out observer-effects, but a future method might aim to minimise observer-effects further.

6.23 The findings from this research suggested that teachers in secondary schools rarely use public criticism of students' work ('TPC') and public positive recognition of their social behaviour ('SBP') - probably for good reason. Given the significant number of zeros accrued in respect of the tallies of 'TPC' and 'SBP', it would be important to consider a re-categorisation of feedback types or redefine category descriptions. There might be a number of novel ways in which teachers in secondary schools provide feedback to students that have not been previously considered that could be incorporated into *MICRO*. Further investigation of this might involve a survey / audit of both students and teachers examining their views about what they think works to increase students' work rates.

6.24 A significant shortcoming of the *MICRO* tool that was previously identified is the temporal disconnect between the recording of an observation and the behaviour that is being observed that is exacerbated by the split-time period observation cycles used. This means that the mechanical connection between behavioural event and a possible behavioural or environmental cause (e.g. an interaction between a particular pupil behaviour and teacher feedback about that behaviour) cannot be examined with any precision. The possibility of using a digital video recording of an observed lesson to cross-reference temporally the observer record could be explored further but this would incur the cost of *MICRO* developing into a less portable and easy-to-use system. Whilst a camera would cause another type of observer-effect requiring measurement and monitoring; and additional ethical considerations and permissions from students and their parents / guardians, the potential of such recordings for future research would be significant however.

The possibility of increasing the representativeness of the sample of observed students in each class could be considered and a modification of Jolly and McNamara's (1992) method incorporated but there has been nothing found in the current study that suggests that a randomly selected sample of 6 students is too small to be representative of all students in a class's behaviour in respect of following the teachers' directions.

Another possibility would be to audio-record and transcribe a lesson in addition to using an updated *MICRO* tool. The possibility of cross-referencing audio-recorded interactions

between teachers and students against a *MICRO* record of a lesson – a development partially explored by Swinson and Harrop (2012; Chapter 2) - seems to offer a way forward.

6.25 A fuller and more detailed quantitative record of interactions in a lesson would provide an opportunity of examining more closely the bi-directional causal effects of verbal interactions in lessons, providing a future version of *MICRO* were enabled to do this.

As has been shown in the literature review (Chapter 2) most of the cited teacher-feedback research projects (White, 1975; Persons, *et al.*, 1976; Nafpaktitus, *et al.*, 1985; Wheldall, *et al.*, 1989; Harrop and Swinson, 2000; Swinson and Harrop, 2001; Apter, *et al.* 2010) that have employed structured classroom observational schedules that predicated the development of the *MICRO* schedule (Appendix) made use of the behaviourist paradigm (Watson, 1913; Pavlov, 1928). In essence, this paradigm depends on a simple concept – a unidirectional association: the behavioural stimulus leads to a behavioural response. Thus feedback-researchers who were interested in teacher-feedback and its effect on students' learning-behaviour chose to concentrate on teacher-feedback as the 'stimulus' in the dyad, and student 'on-task' behaviour as the 'response'.

A unidirectional causal vector was implied, thus:

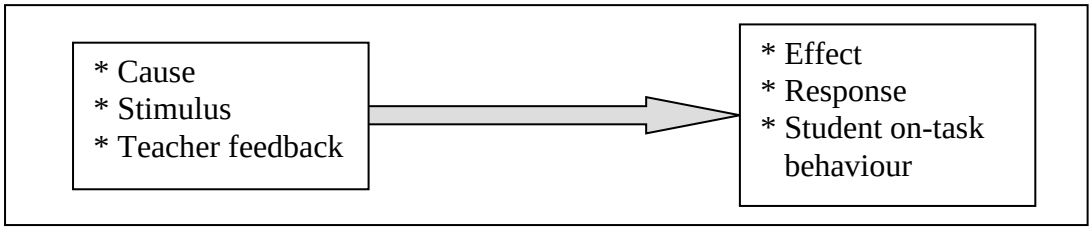


Figure 52 – simple behavioural paradigm applied to teachers' verbal feedback

With Social Learning Theory, Bandura (1977) proposed a more complex relationship between human behaviour and its causes:

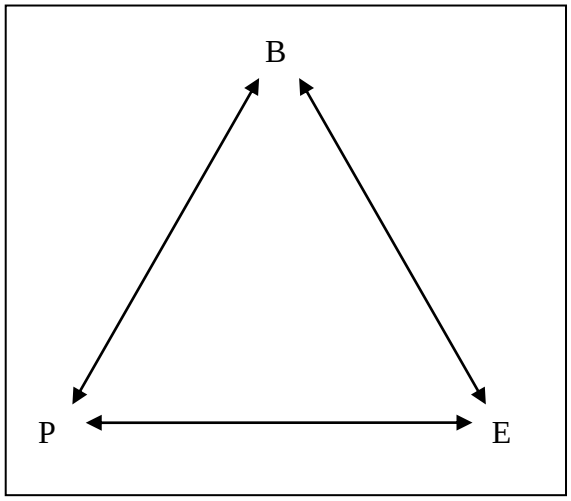


Figure 53 – Bandura's (1986, p.24) diagram of reciprocal determinism

With this model the person and their cognitions (P) have a 2-way causal relationship with their own behaviour (B), and a 2-way causal relationship with the environment (E). The person's behaviour (B) also has a 2-way causal relationship with the environment (E). It is important to note that the environment (E) contains other people - each with their own reciprocally determined behaviour.

Particularly important in this model is the bi-directionality of the causality vectors between each vertex. Statistically, any associations made between activity at a pair of, or all three vertices, is perhaps more accommodatingly and safely expressed therefore with correlations with no implication of causality, and not by multiple regression and the beta statistic.

Bandura developed the temporal dimension of the model in another diagram:

A_1	B_1	A_2	B_2	A_3
S^t ⟷	R ⟷	S^{reinf}		

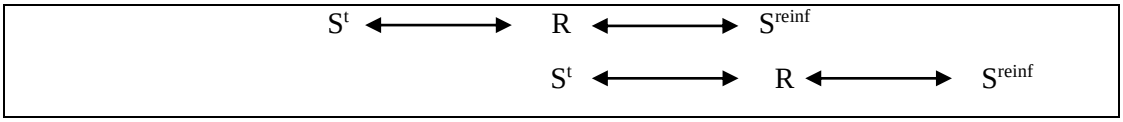


Figure 54 – Bandura's conception of how behaviours reciprocally affect each other

Bandura (1986) explains in a footnote:

> Illustration of how the same events change their status from behavior to environment and from environment to behavior at different entry points in the flow of interaction between two people. The As are successive responses by one person, and the Bs successive responses by the second person in the dual interaction: S^t represents "stimulus"; R represents "response"; S^{reinf} represents "reinforcer." (p.27[50])

In the text Bandura (ibid.) expands on this:

> Confining analysis to a particular interactive segment sheds some light on causal processes. But it inevitably leaves unexplained some of the observed variance in events when other determinants in the triadic system make causal contributions at various points in the transactions. (p.26)

The dynamically-possible complexities that this apparently simple model unleashes over a given time period are astronomical – particularly considering the behavioural relationships of a class of 25 students with their teacher in an academic class in a secondary school.

Developing Bandura's (1986) Social-Cognitive Theory as the underlying paradigm for the results presented here in preference over the behaviourist paradigm means that scepticism about a unidirectional *causal* relationship between teacher's verbal feedback and students' following their directions ('On-task%') is inevitable. Bandura's model indicates a much more complex matrix of social causes and effects operating to promote learning in the classroom[51] than is provided by the findings presented here (Chapters 4 and 5).

However, it is plausible in consideration of the reported findings that the comparatively simplistic correlational association between teachers' verbal feedback ('TPP'; 'SBP'; 'TPC'; and 'SBC') and students' following their directions ('On-task%') has sufficient utility as a *reductive* indication of the way that the bidirectional interaction of teacher-feedback and students' following teachers' directions is contributing to, and quantitatively an indices of, the social-emotional climate for learning. This proposal could be claimed to

[50] Bandura uses the same diagram in 'Social Learning Theory' (1977) but does not use speech marks around 'stimulus', 'response' and 'reinforcer'. Bandura uses speech marks in 1986, because by then he had begun to put a theoretical distance between the triadic reciprocal determinism used by his own social cognitive theory; and stimulus / response determinism used by behaviourism (Watson, 1913; Pavlov, 1928).

[51] The author's updated model of reciprocal determinism (Apter, 2007) is appended (Appendix) and provides a more complex account of the causality vectors that operate within Bandura's model, and which apply to teacher-feedback effects in the classroom.

maintain an indicative compliance with Bandura's conceptual framework of triadic reciprocal determinism (ibid.) without being as exhaustively analytical as it could be.

6.26 The use of MANOVA to analyse results (Chapter 4 and 5) allowed for a more complex interaction of possible associated effects to be examined in deference to Bandura's reciprocal determinism paradigm. These analyses enriched the initial findings of the HMRA analyses that were completed. Grouping together three variables allowed for the bi-directionality of causal vectors between them, and treated student on-task behaviour and teacher verbal feedback behaviour as one entity.

With the UK27 dataset, MANOVA was used to group together 'On-task%'; 'TPP' and 'SBC' as one dependent variable: an amalgamated quantitative indicator of the social emotional climate for learning in the classroom perhaps, and the effect of 'Year group' on that amalgamated measure examined, and the significance of that association established (Chapter 4). A similar significance was established in respect of the association between 'Year group' and the same amalgamated dependent variable within the UK1 dataset (Chapter 5).

In response to the concerns voiced by the SLT of the UK1 school about the differences between maths and English lessons, a second MANOVA was completed in respect of the UK1 dataset that grouped the same three variables together ('On-task%'; 'TPP' and 'SBC') and examined the association with 'Subject' as a single independent variable. Significant differences were established between curriculum subjects (Chapter 5).

6.27 The use of statistical techniques that amalgamate variables such as MANOVA does not however provide detail about bi-directional causality effects – it includes these effects, but opaquely. An alternative way of modelling causality statistically that would have increased compliance with Bandura's reciprocal determinism (1986) would be by using a times-series analysis (Gottman, 1981) but this would require a different temporal organisation of the *MICRO* tool, where simultaneous behavioural events were simultaneously recorded against a timed index.

Sequential analysis incorporating a time-series with an embedded or hidden Markov chain to predict future behaviours (Bart. *et al.*, 1998; Ghahramani, 2001; Zucchini, Macdonald and Langrock, 2016) offer a way forward. Bart. *et al.* explains:

> The basic goal of sequential analysis is to reduce uncertainty about the value of observation by using information about past observations… Sequential analysis is somewhat similar to regression in that an external variable, e.g. time, is used to help understand or predict values of a variable/ (p. 251-252)

And more precisely (Yu, 2016):

> A hidden semi-Markov model is a statistical model. In this model, an observation sequence is assumed to be governed by an underlying semi-Markov process with unobserved (hidden) states. Each hidden state has a generally distributed duration, which is associated with a number of observations produced whilst in the state, and a probability distribution over the possible observations. (Preface)

This statistical method has considerable potential for future investigations involving systematic quantitative observations of classroom behaviour and their analysis.

6.28 Research Question #4

Research question #4 asked: **'**Does the length of periods of time students spend following a teacher's directions in classrooms as measured by *MICRO* ('On-task%'); and a quantitative association with teachers' verbal feedback behaviour, and with other non-verbal, contextual factors, have utility for the senior leadership teams in secondary schools?'

The findings from the ILCS analysis (Apter, 2014) of the senior leadership team (SLT) meeting summarised here (Chapter 5; Appendix) that was convened to discuss the current research, and the subsequent feedback about that meeting, indicated that the presentation of statistical results and draft findings were considered to be of utility for the SLT of the UK1 school. This utility was enhanced by assertive scrutiny, and challenge of both the results and method by a number of members of the SLT team.

6.29 Obviously, the exercise of collecting and analysing quantitative data from the current research might have been more useful to the UK1 school had it been completed prior to the Ofsted inspection in December, 2013, that resulted in the school failing the inspection and being put in to 'special measures'.

The inspection acknowledged that overall, attainments had risen during the previous 12 months but that insufficient 'progress' had been made in English, science, and particularly,

mathematics. The degree of this lack of progress was described in terms of a number of qualitative comments about 'unsatisfactory' lessons that had been observed by Ofsted inspectors but judgements appeared subjective. Whilst the 'quality of teaching' and 'achievement of pupils' was judged inadequate, the observational method was based exclusively upon qualitative guidance and descriptors. Numerical measures were not reported. A quantitative observation method for teachers and lessons was not used. This adheres to published guidance and practice for Ofsted inspections and inspection reports (Ofsted, 2015; see also Ofsted website where inspection reports are publically accessible).

The Ofsted report in 2014[52] following the inspection, judged that the 'behaviour and safety' of pupils was 'good' and students of all ages told inspectors that behaviour in lessons was 'good' and that their learning was 'rarely disrupted'[53]. The previous Ofsted inspection in 2010 had judged the school 'good' in all respects.

The request from the Head Teacher of the UK1 school for the set of observations reported here by the author came about because the UK1 school wanted to understand more comprehensively and more accurately what was required in order to improve the way in which teachers used verbal feedback to students.

Ofsted's judgement about the shortcomings of the maths department and maths lessons in the UK1 school was predicted by the SLT prior to the Ofsted inspection (by all accounts – see Chapter 5). The school's Head Teacher and SLT speculated that if the school had been given access to the findings of the research prior to the Ofsted inspection, they might have been able to hold the inspectors at bay with an action-plan based on the measured baseline provided here presented alongside their maths results.

More importantly and critically, if Ofsted used a more structured and quantitative approach to classroom observation similar to the *MICRO* method demonstrated here with the findings of the current research, their judgements would be – in the author's opinion - fairer and more equitable; more instructive and solution-focussed; less likely to be subverted by political purposes (e.g., pressure to convert to academy status); and ultimately, more respected - rather than feared (and damned) - by schools such as the UK1 school.

[52] Ofsted reports about UK1 school are not referenced in order to protect the school's anonymity.
[53] It could be argued that the absence of a significant difference between the mean 'On-task%' of the UK27 and UK1 datasets reflects this judgement.

6.30 Research question #5

Research question #5 asked: 'Are the length of periods of time students spend following a teacher's directions in classrooms as measured by *MICRO* ('On-task%'); and the positive and encouraging verbal behaviour of teachers reasonable indicators of the conduciveness of the social-emotional climates of classrooms to student learning in UK secondary schools?'

The construct that has been termed for the purposes of the current research: the 'social-emotional climate for learning in classrooms', has not been exhaustively specified in this research, beyond being associated (Chapter 2) with Anderson and Brewer's (1946a,1946b) and Withall's (1949) recognition and elevation of the social-emotional aspects of the classroom that might be more or less conducive to students' learning.

It is suggested here that the association of positive teacher verbal-feedback with students following teachers' directions ('On-task%') is an indicator of the conduciveness of that climate for learning - though initially, there might appear to be a tautological circularity to the relationship between 'indicator', 'climate' and the conduciveness of that climate to learning.

6.31 However, the teacher's behaviour is rather obviously a critical influence in the classroom. Bandura (1977) developed the concept of 'modelling behaviour' whereby he proposed that children were destined to copy adult behaviour from watching them. This mechanism was famously demonstrated in the Bobo Doll experiment (ibid.).
It would seem - however tautological the argument appears - that the social-emotional climate for learning is most effectively measured and developed using structured lesson observations and indicative measures, for example: using *MICRO* or an alternative systematic measure. This would seem to be a particularly ecologically valid proposition in the case of *MICRO*, given that *MICRO* is measuring teaching interactions, teacher-feedback and whether students are following their teachers' directions or not - all artefacts that relate directly to the purposes of a school, whilst at the same time having face-validity as measuring the interactional, social-emotional tone of the classroom[54].

[54] An omission and useful future addition to the *MICRO* schedule might be another of Anderson's innovations, a means of ensuring that teachers reflect on their part in classroom interactions with students: a 'Mental Hygiene Scale For Teachers' (Anderson, 1940)

6.32 The evaluation of the governmental / DfE initiative (Humphrey, Lendrum, and Wiglesworth, 2010) 'Social and Emotional Aspects of Learning (SEAL)' programme in secondary schools described a project that had laudable aims in enhancing students' mental health, resiliency and self-esteem, but which only achieved limited success in terms of outcomes. In common with the current research, Anderson and Brewer's (1946) research and others (Chapter 2), the SEAL programme was based upon a number of untried assumptions. The untried assumptions included: that learning outcomes would be improved by enhancing the emotional well-being of students; and that mental health and resiliency would be improved by manualised, classroom-based psycho-educational activities.

A further obstacle to the success of the SEAL programme was that the mental health and well-being of students was to be primarily measured before and after programme had been implemented principally by the 'Strengths and Difficulties Questionnaire - SDQ' (Goodman, Lamping and Ploubidis, 2010) which was never designed for such a purpose and which - because of its brief and economical structure (5 subscales of 5 items each) - was not a reliable or contextually immune 'test, re-test' measure.

The SEAL programme finally drowned in the wake of an unforgiving quantitative given. The 'given' that would ultimately swamp the programme was that academic outcomes were compelled to become an over-whelming priority in secondary schools in respect of resource allocation - in comparison to any resources previously ring-fenced for student social development, health and well-being (Gove, 2009)[55].

6.33 Research question #6

Research question #6 asked: **'**What is the extent to which secondary school teachers might expect their students to follow their directions in academic lessons (as indicated by the 'On-task%' variable) by year group and as a total-mean of the secondary sample?' The figure obtained by the study can be used as a baseline figure for future secondary research, and so that the teachers are able to make approximate comparisons with classes that they are teaching.

[55] Michael Gove MP who was to become the Secretary of State for Education from May, 2010, until July, 2014, narrowed down the aims of UK education to an academic subset in a speech to the Royal Society for the Arts: 'What Is Education For?', 30th June, 2009, immediately before the publication of the SEAL evaluation in 2010 and his ministerial promotion in the new coalition government.

The mean percentage of time in academic lesson in secondary school classrooms that students were following the directions of their teachers was 78.15% (N= 106, SD= 13.31) for the UK27 dataset and 77.06% (N= 122, SD= 14.4) for the UK1 dataset. The range for the UK27 dataset was from 39% to 99%. The mean figure for primary schools was 85.23% (N= 144, SD= 12.12; Apter, *et al.* 2010).

As has been commented before, it is surprising maybe that these mean 'On-task%' figures for secondary schools which are consistent with each other are significantly lower than the mean on-task percentages found in primary schools (ibid.) and that the mean percentage decrease in the upper secondary year-groups of the UK27 dataset (Chapter 4). It has been previously reported also that teachers use less verbal feedback, both positive and negative, as students grow older. It is of course a school's senior leadership team's decision as to the level of engagement with academic work tasks; and the level of compliance with teachers' directions that teachers in school might reasonably expect; and whether a mean figure of 78.15% is a good enough aim-point, year-group by year-group, and subject by subject, for their secondary school, in comparison to the 85% mean that primary schools appear to achieve.

Chapter 7 – Conclusions

7.0 Response to Introduction (Chapter 1): the role and attributed importance of teachers' verbal feedback in modern secondary school classrooms

7.1 What is feedback and does it work? Hattie (2009), and the Sutton Trust (Higgins, *et al.*, 2014) were adamant about the value of teachers using verbal feedback in the classroom as an integral part of their teaching in order to encourage students to learn. Notwithstanding, Hattie's unusual reworking of what constitutes feedback (Chapter 1), findings from the current research suggest that the relationship between feedback and students engagement with learning in secondary schools in the UK is complex and feedback might not work in quite the same way as it does in primary schools (Chapters 4 and 5).

Findings from the current research suggest strongly that students of different ages receive and internalise teacher-feedback in different ways and that teachers are aware of this and change the way that they deliver feedback, depending upon the age or developmental stage of the students that they are teaching. It is also probable that the longer the teacher has been teaching, the more effective they are in making this adjustment in delivering feedback in an effective age-appropriate way (Chapter 6).

7.2 A number of future investigations have been sign-posted as potentially worthy endeavours by the findings of the current research (Chapter 6).

Whilst an investigation of Dweck's (1999) challenge: *'Caution – praise can be dangerous'* was outside the scope of the current research (Chapter 1), a future direction for research would undoubtedly be to cross-reference the findings presented here with a teachers' 'attitudes and beliefs about feedback' survey; and an investigation of students' phenomenological experiences of teacher-feedback, with culturally contextualised UK samples.

The amount of talking a teacher uses when teaching in secondary schools appears to account for little variation in the engagement of students in academic lessons. A phenomenological survey of secondary students' experiences of teacher-talk and teachers who talk more (or less) would be useful.

Why previous researchers (Chapter 2) found that secondary teachers used more feedback than was observed to be used by teachers in the current research, will probably have to remain un-investigated because of the difficulties of procuring access to their raw data.

Similarly, the correlational association between the percentage of free school meals in a school and the mean on-task rate will probably remain a mystery!

It was a pleasing research finding to be able to report that the more experienced a teacher was, the more engaged and on-task their students were during academic lessons; and also that experienced teachers did not need to use as much criticism of social behaviour as less experienced teachers. It would be interesting to find out what it is that more experienced teachers do differently in the classroom.

A teacher survey would be required to investigate why secondary teachers under-use social behaviour criticism with older students when this type of feedback (more than other types of feedback) appears to work well. Are secondary teachers more anxious about telling older students off when they are disengaged and off-task?

7.3 Social Learning and Social Cognitive Theory and New Reciprocal Determinism

It is proposed that Bandura's Social Learning Theory (1977) and reciprocal determinism (1986) incorporating the author's development (Apter, 2007; Appendix) provides a comprehensive paradigm by which the bi-directional causal vectors of the interactions in classrooms in which students and teachers are involved can be best described and understood.

As has been shown (Chapter 6), Social Learning Theory was a development of behavioural theory (Watson, 1913; Skinner, 1938) and has many features in common, but a critical point of departure is the inclusion in the paradigm of the internal motivational forces within human beings, and of the bi-directional causal vectors of reciprocal determinism. It is suggested that the social-emotional climate for learning in a classroom is better described and examined using reciprocal determinism to model a subset of the quantifiable associations between variables, than was the 'stimulus' and 'response' dyad of behavioural theory.

Acceptance of these corollaries would support the proposal that whilst all the nuances and subtleties of reciprocal determinism will *not* be revealed by the simple indications provided by the version of *MICRO* used for the current research and the limited statistical analysis that has been provided, those indications will not be in any way paradoxical or contradictory to the Social Learning Theory paradigm and triadic bi-directional reciprocal determinism.

7.4 The importance of systematic classroom observation in the applied psychology of Educational Psychologists

It is a conclusion of this secondary study, and of the author's previous primary study (Apter, *et al.* 2010) that Educational Psychologists should persist with using and developing systematic quantitative classroom observation skills as a keystone of their professional practice in schools. Without such a systematic-observation skills-set and associated data-analysis experience, the authority of the applied-psychology of Educational Psychologists in the classroom – working either organisationally or systemically with teachers and school leaders, or when observing individual students as case-work - is diminished.

Other observers of teachers' teaching and students' learning in the classroom such as Ofsted inspectors and teacher-advisors have dispensed with rigorous quantitative observational methods, but this study has shown that data of this type – despite its complexity (or maybe because of it) - has considerable utility in providing provocative and challenging indications of the social emotional climate for learning in classrooms to teachers, senior leaders, managers in schools and policy makers. It is therefore difficult to understand why Ofsted has ceased to aspire to the apolitical objectivity afforded by systematic observation methods of the type described in this study.

Whilst it is clear with the current research that the present state of the technology and statistical analysis used has succeeded and failed[56] in equal measure (Chapter 6) to show how (or indeed, whether) teacher-feedback works in the secondary classroom to optimise students' learning in the classroom, the current un-reductive findings reported here have had

[56] A new *MICRO* schedule is under development. *MICRO* Mk.2 will be less sensitive to observer effects and will facilitate continuous recording of interactional events in the classroom using additional technology. It will also use the knowledge and experience of both primary and secondary teachers in developing a revised, age-sensitive, categorisation of feedback types. The *MICRO* Mk.2 schedule is designed with the facility to use it with a time-series analysis (Gottman, 1981; Bart, Fligner and Notz, 1998) and a hidden semi-Markov statistical model (ibid.; Yu, 2016) in order to investigate the causal relationship between teacher-feedback and student behaviour in a more convincing way than is currently determined by statistical correlation and regression.

an evident utility in inspiring learning conversations within a school leadership team (Chapter 5).

It is the author's opinion that if Educational Psychologists restrict themselves to worthy but small scale qualitative studies, the research of the profession will continue to be sidelined by educational leaders and policy makers who are more influenced and persuaded by the large scale research of educational academics such as Hattie (2009).

7.5 The societal good of a holistically conducive social-emotional climate for learning in schools

The societal good of a holistically conducive social-emotional climate for learning requires more research, discussion and development. The current research illuminates the corollary that was also central to the SEAL programme (Humphrey, *et al.,* 2010): that an improvement in the social-emotional climate for learning in lessons in a school will lead to the school's purposes being more effectively enacted.

The successful meeting of those purposes is not only to be measured by the students' academic attainments however; but by the measured development of every attribute of their presence in the world, including: their knowledge; their curiosity and their motivation to learn; their potential and realised motivation for pro-social action and active citizenship; their physical and emotional well-being; and their political knowledge, ethical judgement skills; and recognition of their familial, community and societal responsibilities.

This is what is promoted in this current research through the adoption of the educationally desirable concept of a holistically conducive social-emotional climate for learning in schools; much as it was promoted by Anderson, and Brewer, (1946); and Withall (1949); and by Smuts, (1926); and by Dewey (1897).

References

Akin-Little, K.A., Eckert, T.L., Lovett, B.J. & Little, S.G. (2004). Extrinsic Reinforcement in the Classroom: Bribery or Best Practice. *Social Psychology Review,* 33, 3, pp. 344 – 362.

Anderson, H.H., (1940). The Construction of a Mental Hygiene Scale for Teachers. *American Journal of Orthopsychiatry,* 10, 2, pp. 253 – 263.

Anderson, H.H. & Brewer, H.M. (1945). Studies of Teachers Classroom Personalities – Dominative and Integrative Behavior of Kindergarten Teachers. *Applied Psychology Monograph,* 6, American Psychological Association / Stanford University Press.

Anderson, H.H. & Brewer, H.M. (1946a). Studies of Teachers Classroom Personalities – Effects of Teachers Dominative and Integrative Contacts on Children's Classroom Behavior. *Applied Psychology Monograph,* 8, American Psychological Association / Stanford University Press.

Anderson, H.H. & Brewer, H.M. (1946b). Studies of Teachers Classroom Personalities – Follow-up Studies of the Effects of Teachers' Dominative and Integrative Contacts on Children's Classroom Behavior. *Applied Psychology Monograph*, 11, American Psychological Association / Stanford University Press.

Apter, B. (2004). TaMBiS Time Sample Record (TSR). *Teaching and Managing Behaviour in Schools.* TaMBiS Course materials for 15 week M.Ed. in-service training module. UK: Wolverhampton City Council / University of Wolverhampton.

Apter, B. (2007). New Reciprocal Determinism. *Teaching and Managing Behaviour in Schools.* TaMBiS Course materials for 15 week M.Ed. in-service training module. UK: Wolverhampton City Council / University of Wolverhampton.

Apter, B., Arnold, C. & Swinson, J. (2010). A mass observation study of student and teacher behaviour in British primary classrooms. *Educational Psychology in Practice,* 26, 2, pp. 151-171.

Apter, B. (2014). Foucauldian Iterative Learning Conversations - an example of organisational change: developing conjoint-work between EPs and Social Workers. *Educational Psychology in Practice,* 30, 4, pp. 331-346.

Association of Teachers and Lecturers – ATL (2013). Schools in special measures in England and Wales. *ATL Facts Sheet.* Retrieved November, 20, 2016, from https://www.atl.org.uk/Images/adv36-schools-in-special-measures-in-england-and-wales-nov-13.pdf

Austin, J. L. (1962). *How to do things with words.* US: Harvard University Press.

Bagley, W. & MacDonald, M. (1932). *Standard Practices in Teaching: a summary of the standards generally recognized as governing good practice in typical teaching procedures.* US: Macmillan.

Baguley, T. (2010). When correlations go bad. *The Psychologist*, 23, 2, pp. 122-123.

Bajpai, N. (2010). *Business Statistics.* India: Pearson Education Inc.

Bakeman, R. & Gottman, J.M. (1997). *Observing interaction: An introduction to sequential analysis*. UK: Cambridge University Press. 2nd edition.

Bandura (1977). *Social Learning Theory*. US: Prentice-Hall.

Bandura, A. (1986). *Social Foundations of Thought and Action.* US: Prentice Hall.

Barr, A.S. (1929). *Characteristic differences in the teaching performance of good and poor teachers of social studies*. US: Public School Publishing Co.

Bart, J., Fligner, M.A. & Notz, W.I. (1998). *Sampling and Statistical Methods for Behavioral Ecologists.* UK: Cambridge University Press.

Berger, P. & Luckmann, T. (1966). *The Social Construction of Reality.* UK: Penguin

Binet, A. & Vaschide, E. (1897). Experiences de forces musculaire et de fond chez les jeunes garcons. *L'année Psychologique,* 4, pp.15-63.

Binet, A. & Simon, T. (1916). *The Development of Intelligence in Children.* (E. Kite, Trans.) US: Williams and Wilkins Co. Baltimore.

Boyle, F. (2009). *My shit life so far.* UK: Harper Collins.

Brenner, B. (1934). Effect of immediate and delayed praise and blame upon learning and recall. *Teachers College Contributions to Education, No. 620.* Retrieved August 15, 2015: http://psycnet.apa.org/index.cfm?fa=search.displayRecord&UID=1934-05405-001

Brooke, S., Tsuei, L. & Malinow, A., (2003). *Thinking Outside of the Box: 100 Years of Educational Psychology at TC.* Retrieved September 4, 2016 from Teacher's College, Columbia University Website: http://www.tc.columbia.edu/news.htm?articleId=4530

Briggs, N. (1935). Practices of best high school teachers. *School Review*, 43, December, pp.745-752.

Bruner, J. (1996). *The Culture of Education*. US: Harvard University Press.

Burgess, D.S. (1981). Book Review: Fifteen Thousand Hours: Secondary Schools And Their Effects On Children. *McGill Journal of Education*. Canada: McGill University

Burt, C.L. (1917). *The Distribution and Relations of Educational Abilities*. UK: The Campfield Press.

Canter, L. & Canter, M. (1976). *Assertive Discipline.* US: Canter and Associates.

Chapman, C.C. & Feder, R.B. (1917). The effect of external incentives on improvement .*Journal of Educational Psychology,* 8, pp.469-474.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 1, pp. 37–46.

Colvin, S.S. (1919). The most common faults of beginning high school teachers. *18th Yearbook - The Professional Preparation of High School Teachers, National Society for the Study of Education, 18, 1.* US: Public School Publishing Co.

Colvin, S.S. (1931). *An introduction to high school teaching.* US: Macmillan.

Dale, E. & Raths, L. (1945). Discussion in the secondary school. *Educational Research Bulletin,* 24, pp. 1-26.

Department for Education - DfE, (2011). Class Size and education in England evidence report. *Research Report. DFE-RR169.* Retrieved August, 20, 2016 from GOV.UK website: www.gov.uk/government/uploads/system/uploads/attachment_data/file/183364/DFE-RR169.pdf

Department for Education - DfE, (2014). Schools, pupils and their characteristics: January 2014. *National Statistics.* Retrieved August, 20, 2016 from GOV.UK website: www.gov.uk/government/statistics/schools-pupils-and-their-characteristics-january-2014

Department for Education - DfE, (2015). Up to 1,000 failing schools to be transformed under new measures - Press release. Retrieved August, 20, 2016 from GOV.UK website: https://www.gov.uk/government/news/up-to-1000-failing-schools-to-be-transformed-under-new-measures

Department for Education - DfE (2015b). National Statistics School workforce in England: November 2015 *National Statistics.* Retrieved July, 10, 2016 from GOV.UK website: https://www.gov.uk/government/statistics/school-workforce-in-england-november-2015

Department for Education - DfE (2015c). Schools, pupils and their characteristics: January 2015. *Statistical First Release – National Statistics.* Retrieved July, 11, 2016 from GOV.UK website: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/433680/SFR16_2015_Main_Text.pdf

Dewey, J. (1897). *My pedagogical creed.* US: Progressive Education Association.

Dewey, J. (1902). *The Child and the Curriculum.* US: University of Chicago Press.

Draper, S.W. (2014). *The Hawthorne, Pygmalion, Placebo and other effects of expectation.* Retrieved August, 28, 2016 from Dept. Psychology, University of Glasgow website: www.psy.gla.ac.uk/~steve/hawth.html

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist,* 41, pp. 1040–1048.

Dweck, C.S. (1999). Caution – Praise Can Be Dangerous. *American Educator,* Spring. US: American Federation of Teachers

Dweck, C.S. (2015). The Secret to Raising Smart Kids. *Scientific American,* January, 23, pp. 76-83.

Elliott, J. (1991). *Action Research for Educational Change.* UK: Open University Press.

Farrell.C. (2015). United Kingdom: School CP, World Corporal Punishment Research. Retrieved January, 29, 2016 from UK website: http://www.corpun.com/counuks.htm

Field, A. (2009). *Discovering Statistics Using SPSS: Third Edition.* UK: Sage.

Flanders, N.A. (1964). Some relationships among teacher influence, pupil attitudes and achievement. In Biddle, B.J. & Ellena, W.J. [Eds.] *Contemporary Research on Teacher Effectiveness.* US: Holt, Rinehart and Wilson.

Foucault, M. (1969). *L'archéologie du savoir.* France: Gallimard.

Galton, M. (1978). Systematic Classroom Observation: British Research. *Educational Research,* 21, 2 pp. 109-115.

Gilbreth, L. M. (1917). *The Psychology of Management: the Function of the Mind in Determining, Teaching and Installing Methods of Least Waste*. New York: Sturgis and Walton.

Gottman, J.M. (1981). *Time-Series Analysis: A Comprehensive Guide for Social Scientists*. UK: Cambridge University Press.

Güner Yildiz, N. & Sazak Pinar, E. (2014). Examining Approval and Disapproval of Teachers Working in Inclusive Classrooms. *International Journal of Instruction*, 7, 1, pp. 33 – 46.

Ghahramani, Z. (2001). An Introduction to Hidden Markov Models and Bayesian Networks. *International Journal of Pattern Recognition and Artificial Intelligence,* 15, 1, pp. 9 – 42.

Gilchrist, E.P. (1916). The extent to which praise and reproof affect a pupil's work. *School and Society,* 4, 101, pp.872-874.

Goodenough, F.L. (1928). Measuring behavior traits by means of repeated short-sample. *Journal of Juvenile Research,* 12, pp. 230-235.

Goodenough, F.L. (1930). Inter-Relationships in the Behavior of Young Children. *Child Development,* 1, 1, pp. 29-47.

Goodenough, F.L. (1937). The Observation of Children's Behaviors as a Method in Social Psychology. *Social Forces*, 15, 4, pp. 476-479.

Goodman, A., Lamping, D.L. & Ploubidis, G.B. (2010). When to use broader internalising and externalising subscales instead of the hypothesised five subscales on the Strengths and Difficulties Questionnaire (SDQ): data from British parents, teachers and children. *Journal of Abnormal Child Psychology,* 38, 1179-1191.

Gottman, J.M., (1981). *Time-Series Analysis – A Comprhensive Introduction for Social Scientists.* UK: Cambridge University Press.

Gove, M. (2009). What Is Education For? *Speech to the Royal Society for the Arts, 30 June 2009.* Transcription retrieved May, 20, 2016, from the UK website: https://www.thersa.org/globalassets/pdfs/blogs/gove-speech-to-rsa.pdf

Green, S.B., McCoy, J.F., Burns, K.P. & Smith, A.C. (1982). Accuracy of observational data with whole interval, partial interval and momentary time-sampling recording techniques. *Journal of Behavioral Assessment*, 4, pp. 103-118.

Guest, G.S., MacQueen, K.M. & Namey, E.E. (2012*). Applied Thematic Analysis.* UK: Sage.

Harding, J. (2013). *Qualitative Data Analysis from Start to Finish.* UK: SAGE Publishers.

Harris, Z. (1988). *Language and Information.* US: Columbia University Press.

Harrop, A. & Swinson, J. (2000). Natural rates of approval and disapproval in British infant, junior and secondary classrooms. *British Journal of Educational Psychology*, 70, pp. 473 – 483.

Hattie, J.A.C. (2009). *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement.* UK: Routledge.

Heath, A. & Clifford, P. (1980). The Seventy Thousand Hours that Rutter Left Out. *Oxford Review of Education*, 6, 1, pp. 3-19.

Higgins, S., Katsipataki, M., Kokotsaki, D., Coleman, R., Major, L.E. & Coe, R. (2014). *The Sutton Trust-Education Endowment Foundation Teaching and Learning Toolkit.* Retrieved, January, 20, 2016, from the UK Education Endowment Foundation website: https://archive.org/details/intelligencesch00termgoog

HM Government (1987). Statutory Instrument: 1987 No. 461 (N.I. 6), Northern Ireland, The Education (Corporal Punishment) (Northern Ireland) Order 1987. Retrieved, January, 21, 2016, from the UK *National Archives* website: http://www.legislation.gov.uk/nisi/1987/461/introduction/made

Hobbs, G. & Vignoles, A., (2007). Is Free School Meal Status a Valid Proxy for SocioEconomic Status (in Schools Research)? *CEE DP 84 Report.* Retrieved, July 7, 2016, from the Centre for the Economics of Education, London School of Economics website: http://webarchive.nationalarchives.gov.uk/20130401151715/http://www.education.gov.uk/publications/eOrderingDownload/CEE03-07FR.pdf

Hoetker, J. & Ahlbrand, W.P. (1968). The Persistence of the Recitation: A Review of Observational Studies of Teacher Questioning Behaviour. *Occasional Paper Series, No.3*. USA: Central Midwestern Regional Educational Laboratory.

Houghton, S., Wheldall, K., Jukes, R. & Sharpe, A. (1990). The effects of limited private reprimands and increased private praise on classroom behaviour in four British secondary school classes. *British Journal of Educational Psychology,* 60, pp. 255 – 265.

Humphrey, N., Lendrum, A. & Wigelsworth, M. (2010). Social and emotional aspects of learning (SEAL) programme in secondary schools: national evaluation - Research Report DFE-RR049. Retrieved February 10, 2016, from GOV.UK website: https://www.gov.uk/government/uploads/system/uploads/...data/.../DFE-RR049.pdf

Hurlock, E.B. (1924). The value of praise and reproof as incentives for children. *Archives of Psychology,* 71.

Hurlock, E.B. (1925). An evaluation of certain incentives used in school work. *The Journal of Educational Psychology,* 16, 3, March pp. 145-159.

Hurlock, E.B. (1931). The psychology of incentives. *The Journal of Social* Psychology, 2, 3, August, pp. 261-289.

Hymer, B., Whitehead, J., & Huxtable, A.M. (2008). *Gifts, Talents and Education: A Living Theory Approach.* UK: Wiley-Blackwell.

Jewitt, C. (2012). An introduction to using video for research. *National Centre for Research Methods Working Paper.* Retrieved January, 25, 2016, from Institute of Education website: eprints.ncrm.ac.uk/2259/4/NCRM_workingpaper_0312.pdf

Johannesson, I. (1967). The effects of praise and blame. *Stockholm Studies in Educational Psychology,* 13. Stockholm: Almqvist & Wiksell.

Jolly, M. & McNamara, E. (1992). *Pupil Behaviour Schedule.* Retrieved May, 4, 2008, from UK: Positive Behaviour Management website: www.positivebehaviourmanagement.co.uk.

Joncich, G. (1968). *The Sane Positivist.* US: Wesleyan University Press.

Jones, L. (1924). *The Training of Teachers in England and Wales.* Oxford: OUP.

Kandel, E.R, & Tauc, L. (1965). Heterosynaptic facilitation in neurons of the abdominal ganglion of aplysia depilans. *Journal of Physiology,* 181, pp. 1–27.

Kennedy, W.A., Turner, A.J., Lindner, R. (1962). Effectiveness of praise and blame as a function of intelligence *Perceptual and Motor Skills,* 15, pp. 143 - 149

Kennedy, W.A. & Vega, M. (1965). Negro children's performance on a discrimination task as a function of examiner's race and verbal incentive. *Journal of Personality and Social Psychology,* 2, pp. 839-843.

Kennedy, W.A. & Willcutt, H.C. (1964). Praise and blame as incentives. *Psychological Bulletin,* 62, 5, pp. 323-332

Kennedy, W.A. & Willcutt, H.C. (1965). A factorial study of praise and blame *Psychological Reports,* 17, pp. 511-517.

Keating, J. (2010). *Teacher training - up to the 1960s.* Retrieved May 4, 2016, History in Education Project, Institute of Historical Research, University of London website: http://www.history.ac.uk/history-in-education/project-papers/topics

Khoo, C., Na, J-C & Jaidka, K., (2011). Analysis of the macro-level discourse structure of literature reviews. *Online Information Review*, 35, pp. 255-271 Retrieved June 3, 2016 from website: http://www.emeraldinsight.com/doi/abs/10.1108/14684521111128032

Kirby, T.J. (1913). Practice in the case of school children. *Teachers College Contribution to Education, No. 58.* Retrieved November, 15, 2015, from website: https://archive.org/stream/practiceincaseof00kirbrich#page/6/mode/2up

Kounali, D., Robinson, T., Goldstein, H. & Lauder, H. (2014). The probity of free school meals as a proxy measure for disadvantage. *Academia.edu / Department of Education Studies University of Bath.* Retrieved May 3, 2016, from University of Bath website: http://www.bath.ac.uk/research/harps/Resources/The%20probity%20of%20FSM%20revised%2014.7.08.pdf

Landsberger, H.A. (1958). *Hawthorne Revisited.* NY: Ithaca, The New York State School of Industrial and Labour Relations.

Leary, M. (2014). *Classroom Observation – a guide to the effective observation of teaching and learning.* UK: Routledge.

Lewin, K.H.F. (1936). *Principles of topological psychology.* (Heider, G. M. Trans.) US: McGraw-Hill.

Liao, S.C., Hunt, E.A. & Chen, W. (2010). Comparison between Inter-rater Reliability and Inter-rater Agreement in Performance Assessment. *Annals of Academy of Medicine Singapore*, 39, pp. 613 – 618.

Madill, A. & Gough, B. (2008). Research and Its Place in Psychological Science. *Psychological Methods*, 13, 3 pp. 254-271.

McNamara, E. & Jolly, M. (1990). Are Disruptive Behaviours Reduced When Levels of On-task Behaviours Increase? An Across Settings Study of a Class of 12- and 13-Year-Old Pupils. *Behavioural Psychotherapy*, 18, pp. 239-249.

Madsen, C.H., Becker, W.C. & Thomas, D.R. (1968). Rules, Praise and Ignoring: Elements of Elementary Classroom Control. *Journal of Applied Behavioral Analysis*, 1, 2, pp. 139-150.

Mansell, W. (2011). No choice but to become an academy. *The Guardian,* 19.12.11. Retrieved January, 14, 2016, from UK Guardian website: https://www.theguardian.com/education/2011/dec/19/schools-forced-to-become-academies

Mayo, E. (1933). *The human problem of an industrial civilisation*. US: MacMillan.

McBer, H. (2000). *Research into Teacher Effectiveness A Model of Teacher Effectiveness: Research Report No 216.* Retrieved February 4, 2016, from UK Department for Education and Employment website: http://dera.ioe.ac.uk/4566/1/RR216.pdf

McIntyre, D. (1980). Systematic observation of classroom activities. *Educational Analysis*, 2, pp. 3-30.

Merrett, K., & Wheldall, K. (1986). Observing Pupils and Teachers in Classrooms (OPTIC): a behavioural observation schedule for use in schools. *Educational Psychology,* 6, 1, pp. 57-70.

Merrett, F. & Wheldall, K. (1987). Natural rates of teacher approval and disapproval in British primary and middle school classrooms. *British Journal of Educational Psychology,* 57, pp. 95-103.

Merrett, F. & Wheldall, K. (1992). Teachers use of praise and reprimands to boys and girls. *Educational Review,* 44, 1 pp. 73-80.

Meyer, W.U. (1992). Paradoxical effects of praise and criticism on perceived ability. In W. Strobe & M. Hewstone (Eds.), *European review of social psychology* Vol. 3, pp. 259–283, UK: Wiley.

Milar. C. & Hawkins, R. (1976). Distorted results from the use of interval recording techniques. In T.Brigham, R.Hawkins, J.Scott, T.McLoughlin (Eds.), *Behavior analysis in education: Self-control and reading,* US: Kendall Hunt.

Monroe, W.S. (1928). *The high school - Teacher training series*. Doubleday, Doran & Company, Inc.

Monroe, W.S. & Carter, R.E. (1923). *The use of different types of thought questions in secondary schools and their relative difficulty for students.* Bureau of Educational Research, 14, pp.5-6 US: College of Education, University of Illinois

Nafpaktitus, M., Mayer, G.R. & Butterworth, T. (1985). Natural Rates of Teacher Approval and Disapproval and Their Relation to Student Behavior in Intermediate School Classrooms. *Journal of Educational Psychology,* 77, 3, pp. 362-367.

Nagel, E. (1961). *The Structure of Science*. US: Blackwell.

O'Connell, A.N. (1990). *Women in Psychology: A Bio-bibliographic Sourcebook.* US:Greenwood.

OECD (2012). How does class size vary around the world? *Education Indicators In Focus.* Retrieved May 5, 2016, from OECD website: http://www.oecd.org/edu/skills-beyond-school/EDIF%202012--N9%20FINAL.pdf

Ofsted (1995). *Class size and the quality of education.* UK: Office for Standards in Education.

Ofsted (2015). *School inspection handbook January 2015, No. 120101.* Retrieved April 5, 2016, from UK Ofsted website: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/391531/School_inspection_handbook.pdf

O'Leary, M. (2014). *Classroom Observation - A guide to the effective observation of teaching and observation.* UK: Routledge.

Olson, W.C. (1929). The Measurement of Nervous Habits in Normal Children. *Monograph, 3.* US: University of Minnesota.

Olson, W.C. (1931). A Study of Classroom Behaviour. *Journal of Educational Psychology,* 22, pp. 449-454.

Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science,* 349, 6251, p. 943. Retrieved September 12, 2016, from OSC website: dx.doi.org/10.1126/science.aac4716

Ostrov, J.M. & Hart, E.J. (2013). Observational methods. In *The Oxford Handbook of Quantitative Methods in Psychology, Volume 1,* (Ed. T.D. Little). UK: OUP.

Painter, F.V.N. (1896). *A History of Education.* NY: D. Appleton and Co.

Palint, J. (2010). *SPSS Survival Manual.* UK: McGraw Hill.

Pashler, H. & Wagenmakers, E. (2012). Editors' Introduction to the Special Section of Replicability in Psychological Science: A Crisis of Confidence? *Perspectives on Psychological Science,* 7, 6, pp. 528-530.

Paton, G., (2014). UK class sizes among biggest in the developed world. *The Telegraph,* September 9, 2014. Retrieved from UK Telegraph Newspapers website: http://www.telegraph.co.uk/education/educationnews/11083683/OECD-UK-class-sizes-among-biggest-in-developed-world.html

Pavlov, I. P. (1928). *Lectures in Conditioned Reflexes*. (W.H.Gantt, 1928, Trans.). US: Liveright Publishing Corp.

Persons, W.S., Brassel, W.R. & Rollins, H.A. (1976). A practical observation procedure for monitoring four behaviors relevant to classroom management. *Psychology in the Schools*, 13, pp. 64-71.

Pett, M.A. (2016). *Nonparametric statistics for health care research: statistics for small samples and unusual distributions: 2nd Edition.* US:Sage.

Piaget, J. (1926). *The Language and Thought of the Child.* UK: Routledge & Kegan Paul.

Piaget, J. (1954). *The Construction of Reality in the Child.* US: Basic Books Inc.

Pickren, W.E., Dewsbury, D.A., & Wertheimer, M. (2012). *Portraits of Pioneers in Developmental Psychology*. US: Taylor and Francis.

Plato, (2002). *Five Dialogues: Euthyphro, Apology, Crito, Meno, Phaedo, 2nd Edition* Ed. J.Cooper (G.M.A. Grube, Trans. originally published BCE 390). UK: Hackett Publishing

Popper, K. (1935). *Logik der Forschung.* Austria: Verlag von Julius Springer.

Powell, J. Martindale, B., Kulp, S., Martindale, A. & Bauman, R. (1977). Taking a closer look: time sampling and measurement errors. *Journal of Applied Behavior Analysis,* 10, pp. 325-332.

Price, B. (1989). Frank and Lillian Gilbreth and the Manufacture and Marketing of Motion Study, 1908-1924. *Business and Economic History, Series 2, 18.* US: Business History Conference (proceedings).

Rapp, J.T., Carroll, L.S., Swanson, G. & Higgins, W.J. (2015). A Comparison of Reliability Measures for Continuous and Discontinuous Recording Methods: Inflated Agreement Scores With Partial Interval Recording and Momentary Time Sampling for Duration Events. *Behavior Modifiication,* 35, 4, pp. 389 – 402.

Rich, R W. (1933). *The training of teachers in England and Wales during the nineteenth century*. UK: Cambridge University Press.

Roethlisberger, F.J. & Dickson, W.J. (1939). *Management and the Worker*. US: Harvard University Press.

Rogers, C. (1951). *Client-Centred Therapy,* UK: Constable.

Rosenthal, R. & Jacobson, L. (1968). *Pygmalion in the Classroom: Teacher expectation and pupils' intellectual development.* US: Irvington Publishers.

Rutter, M., Maughan, B., Mortimore, P., Ouston, J. & Smith, A., (1979). *Fifteen Thousand Hours: Secondary Schools and their Effects on Children*. UK: Open Books.

Sazak Pinar, E. & Güner Yildiz, N. (2013). Investigating Teachers' Approval and Disapproval Behaviors towards Academic and Social Behaviors of Students with or without Special Needs. *Educational Science: Theory and Practice*, 13, 1, pp. 551 – 556.

Schmidt, H.O., (1941). The Effects of Praise and Blame as Incentives to Learning. *Psychological Monographs,* 53, 3. US: American Psychological Association.

Segal, S. S. (1968). *No child is ineducable: Special Education: Provision and Trends.* US: Pergamon Press.

Simpson, L. (1980). 15,000 Hours: Summary. *Radical Statistics,* 18. Retrieved from Radical Statistics website: http://www.radstats.org.uk/no018/

Sinclair, J.H. & Coulthard, M. (1974). *Towards an Analysis of Discourse: English Used by Teaches and Pupils.* UK: Oxford University Press.

Skiba, R.J. (1989). Temporal Parameters in the Sampling of Behaviour: The Accuracy and Generalizability of Observation. *Monograph, No.10, Instructional Alternatives Project.* US: University of Minnesota.

Skinner, B.F. (1938). *The Behavior of Organisms: An Experimental Analysis*. US: Cambridge, Mass. B.F. Skinner Foundation.

Smith, J. (2009). *Interpretative Phenomenological Analysis.* UK: Sage.

Smuts, J.C. (1926). *Holism and Evolution.* UK: Macmillan.

Steffen, W.F. (1985). *An empirical study of measurement error using partial interval, whole interval, and momentary time-sample recording methods: thesis*. US: School of Graduate Studies, Drake University.

Steller, A.W. (1980). Review: Fifteen Thousand Hours: Secondary Schools and Their Effects on Children. *Educational Leadership,* October, 1980, p. 90.

Stevens, R. (1912). The Question as a Measure of Efficiency in Instruction. *Teachers College Columbia University – Contributions to Education No. 48, Sixth Impression.* New York City: Teachers College Columbia University.

Strauss, A. & Corbin, J.M. (1990). *Basics of Qualitative Research: Grounded Theory Procedures and Techniques 2nd (second) Edition.* UK: Sage.

Sturt, M, (1967). *The Education of the People.* UK: Routledge and Kegan Paul.

Sutton Trust EEF (2015). *Private tuition polling 2015.* Retrieved May 5, 2016, from the UK Sutton Trust EEF website: http://www.suttontrust.com/researcharchive/private-tuition-polling-2015/

Swinson, J. & Harrop, A. (2001). The Differential Effects of Teacher Approval and Disapproval in Junior and Infant Classrooms. *Educational Psychology in Practice,* 17, 2, pp. 157 – 167.

Swinson, J., (2005). *The quality and quantity of teacher verbal feedback and its relationship to pupil behaviour before and after training.* UK: Unpublished doctoral thesis, John Moores University, Liverpool.

Swinson, J. & Harrop, A. (2012). *Positive Psychology for Teachers*. UK: David Fulton / Routledge.

Tabachnick, B. G. & Fidell, L. S. (2012). *Using Multivariate Statistics (6th ed.)*. US: Pearson.

Terman, L.M. (1916). *The Intelligence Of School Children – How Children Differ in Ability. The use of Mental Tests in School Grading and the Proper Education of Exceptional Children.* US: Houghton Mifflin Co.

Thomas, J.D., Presland, M., Grant, D. & Glynn, T. (1978). Natural rates of teacher approval and disapproval in grade 7 classrooms. *Journal of Applied Behavior Analysis,* 11, 1, pp. 91-94.

Thorndike, E., (1898). Some experiments on animal intelligence. *Science,* 7, 181, pp. 818-824.

Thorndike, E. (1903). *Educational Psychology.* NY: The Science Press

Thorndike, E. (1911). *Animal Intelligence.* NY: Macmillan

Thorndike, E. (1927). The Law of Effect. *American Journal of Psychology,* 39, pp. 212-222.

Tiber, N. & Kennedy, W. A. (1964). The effect of incentives on the intelligence test performance of different social groups. *Journal of Consulting Psychology* , 28, 187, pp. 187-189.

Todes, D.P. (2014). *Ivan Pavlov: A Russian Life in Science.* UK: Oxford University Press.

Turner, C. (2012). The Hidden Persuader. *Cabinet*, 42, pp. 28-32.

Walsall Observer (1975). Discipline suffering at schools. *Walsall Observer*, 27th June.1975. Retrieved May 14, 2016, from Walsall Observer archive website: http://s88.photobucket.com/user/corpuncom/media/archive2/22806a.jpg.html

Warnock, M. (1978). *Special Educational Needs. Report of the Committee of Enquiry into the Education of Handicapped Children and Young People.* UK: HMSO.

Watson, J. B. (1913). Psychology as the Behaviorist Views it. *Psychological Review* 20, pp. 158–177.

Wheldall, K., Merrett, F. & Borg, F. (1985). The Behavioural Approach to Teaching Package (BATPACK): An Experimental Evaluation. *British Journal of Educational Psychology*, 55, pp. 65-75.

Wheldall, K., Houghton, S. & Merrett, F. (1989). Natural rates of teacher approval and disapproval in British secondary classrooms *British Journal of Educational Psychology*, 59, pp. 38-48.

White, M.A. (1975). Natural Rates of Teacher Approval and Disapproval in the Classroom. *Journal of Applied Behavioral Analysis*, 8, 4, pp. 367-372.

Whitehurst, G.J. & Chingos, M.M. (2011). *Report: Class Size: What Research Says and What it Means for State Policy.* US: Brown Centre on Education Policy at Brookings.

Winnicott, D.W. (1976). *The Maturational Processes and the Facilitating Environment*. UK: Hogarth Press.

Withall, J. (1949). The development of a technique for the measurement of social-emotional climate in classrooms. *Journal of Experimental Education*, 17, 3, pp. 347-361.

Wragg. (1993). *Primary Classroom Skills.* UK: Routledge.

Wragg, E.C. (1994). *An introduction to classroom observation.* UK: Routledge.

Wragg, E.C. (1999). *An introduction to classroom observation, second edition.* UK: Routledge.

Xenophon, (1990). *Conversations of Socrates*. Ed. Robin Waterfield (H.Tredennick, Trans. originally published BCE 371). UK: Penguin Classics.

Yu, S-Z., (2016). *Hidden Semi-Markov Models: Theories, Algorithms and Applications.* Netherlands: Elsevier Inc.

Ziliak, S.T. & McCloskey, D.N. (2008). *The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice, and Lives*. US: University of Michigan Press

Zucchini, W., MacDonald, I.L. & Langrock, R. (2016). *Hidden Markov Models for Times Series: An Introduction Using R, Second Edition.* UK: Chapman and Hall.

Appendix – Raw Data and Textual Materials

UK27 and UK1 Raw Data (SPSS output to PDF)

The following pages show 14 displays of raw data from one SPSS data entry sheet that includes both UK27 and UK1 datasets. The datasets have been anonymised so that individual schools and observers cannot be recognised.

KEY

LocType	-	Locality Type: 1. inner city, 2. suburban or 3. rural
FSM	-	Free school meals as a percentage
Roll	-	No. of students
Yr Gp	-	Year group
M_F	-	Gender of teacher
YrsTch	-	Years teaching (experience)
Subj.	-	Subject
Adlts	-	Number of adults in classroom
ClasSz	-	Class size
AMPM	-	AM or PM lesson
OnTsk	-	Time on-task as a percentage
INX	-	Teacher-talking time as a percentage
TPP	-	Task performance positive comment
SBP	-	Social behavioural positive comment
TPC	-	Task performance critical coment
SBC	-	Social behavioural criticism
TME	-	Time: duration of observation
UK1_2	-	Data row: 1. UK1, or 2. UK27

	LocTyp	FSM	Roll	YrGp	M_F	YrsTch	Subj	Adlts	ClasSz	AMPM	OnTsk
1	1	71	861	9	2	1	9	2	15	1	92
2	1	71	861	9	1	8	9	2	18	1	89
3	1	71	861	10	2	12	3	1	8	1	76
4	1	71	861	7	2	5	9	3	16	1	91
5	1	71	861	7	2	17	9	2	16	2	88
6	1	39	1199	7	2	22	9	1	10	1	75
7	1	39	1199	7	2	22	9	4	10	1	81
8	1	17	1251	9	1	2	5	1	28	1	69
9	1	17	1251	11	2	20	9	2	5	2	94
10	2	9	885	9	2	24	3	1	19	1	77
11	2	9	885	9	2	24	3	6	20	2	70
12	2	64	788	7	2	8	9	2	26	1	83
13	2	64	788	7	2	9	9	1	21	2	89
14	2	8	912	9	2	4	1	1	28	1	82
15	2	8	912	8	2	4	1	1	9	2	79
16	1	21	946	7	2	12	1	2	9	1	98
17	1	21	946	8	2	12	1	1	16	2	99
18	2	6	1651	8	1	25	3	3	26	1	91
19	2	6	1651	8	1	25	3	3	24	2	86
20	2	6	824	8	1	20	2	1	19	1	79
21	2	6	824	11	1	20	2	1	23	2	75
22	1	35	1480	8	2	12	2	1	25	1	81
23	1	35	1480	8	2	4	9	1	26	2	88
24	2	5	640	8	2	4	10	1	30	1	84
25	2	5	640	8	2	4	10	1	25	2	78
26	2	11	1486	8	2	15	9	1	26	1	97
27	2	11	1486	9	2	15	9	1	30	1	98
28	1	20	1173	7	1	8	2	2	27	1	95
29	1	20	1173	7	1	3	9	1	7	2	97
30	2	6	940	8	1	8	2	2	27	1	84
31	2	6	940	8	1	8	2	1	27	2	85
32	2	5	831	8	2	3	2	1	27	1	79
33	1	5	831	10	2	3	2	1	29	1	78
34	1	9	931	8	1	6	2	1	14	1	75
35	2	9	931	7	1	1	2	1	22	1	64
36	2	.	1100	8	1	13	2	3	14	1	93
37	2	.	1100	7	1	4	2	3	22	1	76
38	2	18	1049	12	2	10	9	1	28	1	82
39	2	18	1049	12	2	14	9	1	24	2	57

	INX	TPP	SBP	TPC	SBC	TIME	Uk1_2
1	45	34	0	2	14	36	2
2	56	38	0	0	4	36	2
3	83	7	0	0	7	30	2
4	51	32	0	3	7	36	2
5	80	8	0	1	1	30	2
6	91	27	3	3	3	30	2
7	38	22	0	1	8	34	2
8	64	9	0	21	12	30	2
9	43	12	3	24	1	32	2
10	70	20	1	25	12	34	2
11	71	41	6	16	29	36	2
12	82	43	6	13	27	36	2
13	64	9	0	1	28	36	2
14	53	23	0	2	8	36	2
15	62	14	4	9	4	36	2
16	65	22	2	1	2	30	2
17	62	30	1	8	16	30	2
18	66	49	3	6	22	36	2
19	67	24	1	1	25	36	2
20	59	13	0	2	4	36	2
21	57	5	1	10	17	36	2
22	100	0	0	0	0	30	2
23	36	3	0	3	2	36	2
24	21	11	15	6	18	36	2
25	30	2	2	1	19	38	2
26	39	12	2	2	0	30	2
27	26	3	4	2	10	34	2
28	60	1	0	1	3	30	2
29	53	0	0	2	1	30	2
30	53	7	1	0	3	30	2
31	77	16	1	3	7	30	2
32	55	12	1	3	5	30	2
33	39	14	0	2	12	30	2
34	45	15	1	13	3	32	2
35	36	8	0	9	28	32	2
36	56	25	3	16	4	30	2
37	51	17	0	21	39	30	2
38	50	16	14	7	6	36	2
39	43	7	9	2	18	30	2

	LocTyp	FSM	Roll	YrGp	M_F	YrsTch	Subj	Adlts	ClasSz	AMPM	OnTsk
40	2	9	1683	9	2	11	2	1	30	1	95
41	2	9	1683	11	2	11	2	1	29	2	71
42	2	8	1400	8	2	0	9	1	30	1	62
43	2	8	1400	7	2	5	9	1	23	1	75
44	2	8	1400	7	2	5	9	1	29	1	69
45	2	8	1400	8	2	5	1	1	27	2	65
46	2	8	1400	9	2	6	1	2	14	2	72
47	2	8	1400	10	1	7	2	1	23	1	77
48	2	8	1400	7	1	0	2	2	27	1	65
49	2	8	1400	8	2	4	3	1	30	1	54
50	2	8	1400	9	2	3	9	1	12	1	48
51	2	8	1400	8	2	3	9	1	12	2	56
52	1	23	543	7	1	7	3	1	20	1	83
53	1	23	543	7	1	5	3	2	14	1	89
54	1	23	543	8	2	6	1	1	19	1	67
55	1	23	543	11	2	6	1	2	14	2	39
56	1	23	543	9	1	7	9	1	13	1	89
57	1	23	543	11	2	1	9	1	12	1	45
58	1	23	543	7	1	5	3	1	20	1	86
59	1	23	543	10	1	5	3	1	10	2	81
60	1	23	543	11	1	5	3	1	9	1	63
61	1	23	543	7	1	5	3	2	14	1	83
62	1	23	543	11	1	9	5	1	13	1	94
63	1	23	543	10	1	9	5	1	24	2	85
64	1	23	543	7	1	5	3	1	13	1	78
65	1	23	543	10	1	5	3	1	10	2	87
66	1	23	543	9	2	9	1	1	25	1	98
67	1	23	543	9	2	9	1	1	25	1	86
68	1	23	543	7	1	5	3	1	20	1	84
69	1	23	543	10	1	5	3	1	10	2	66
70	1	23	543	9	1	7	9	1	13	1	87
71	1	23	543	11	2	1	9	1	12	1	62
72	1	23	543	7	1	5	3	1	20	1	81
73	1	23	543	7	1	6	3	2	14	2	84
74	1	23	543	8	2	6	1	1	19	1	73
75	1	23	543	11	2	6	1	2	14	2	55
76	2	58	1234	11	2	8	1	2	10	1	89
77	2	58	1234	12	2	7	1	1	4	2	95
78	1	23	543	11	1	9	5	1	13	1	80

	INX	TPP	SBP	TPC	SBC	TIME	Uk1_2
40	45	18	12	3	0	30	2
41	32	15	3	1	1	30	2
42	55	21	5	15	12	30	2
43	54	15	1	3	14	30	2
44	83	18	3	10	14	30	2
45	56	28	3	3	14	32	2
46	65	24	0	36	25	30	2
47	84	23	3	7	21	30	2
48	68	29	2	12	32	30	2
49	71	7	1	5	19	30	2
50	73	17	2	1	29	30	2
51	53	23	1	1	27	30	2
52	28	31	8	2	9	32	2
53	33	32	6	0	3	32	2
54	33	17	12	13	25	38	2
55	22	4	6	1	10	30	2
56	41	16	6	7	22	32	2
57	16	6	0	1	7	32	2
58	35	15	4	0	14	30	2
59	28	6	3	0	4	30	2
60	48	7	1	17	31	32	2
61	38	10	0	0	10	32	2
62	47	8	1	4	15	30	2
63	42	11	0	1	21	30	2
64	19	10	0	8	15	32	2
65	42	3	0	3	7	30	2
66	38	27	0	0	4	30	2
67	38	11	0	1	5	30	2
68	37	34	9	3	10	38	2
69	35	9	.	0	0	38	2
70	24	17	4	10	25	38	2
71	16	6	4	2	3	38	2
72	48	27	8	2	5	38	2
73	48	26	8	0	0	38	2
74	49	21	5	11	31	38	2
75	22	22	1	3	8	34	2
76	59	4	0	1	1	34	2
77	85	1	0	1	1	34	2
78	44	9	0	4	14	36	2

	LocTyp	FSM	Roll	YrGp	M_F	YrsTch	Subj	Adlts	ClasSz	AMPM	OnTsk
79	1	23	543	10	1	9	5	1	24	2	87
80	1	23	543	7	1	15	3	1	13	1	89
81	1	23	543	10	1	15	3	1	14	2	86
82	1	23	543	11	1	7	3	1	9	1	86
83	1	23	543	7	1	5	3	2	14	1	78
84	1	23	543	10	2	9	1	1	24	1	75
85	1	23	543	10	2	9	1	1	24	1	80
86	2	8	1300	8	2	6	2	1	18	1	84
87	2	8	1300	8	2	6	2	1	18	2	87
88	2	8	1300	8	1	8	2	1	19	1	79
89	2	8	1300	11	1	8	2	1	23	2	75
90	2	8	1300	7	1	8	2	1	29	2	71
91	2	8	1300	7	1	7	2	1	30	1	95
92	2	8	1300	11	1	15	2	1	14	1	93
93	2	8	1300	9	1	14	2	1	22	1	76
94	1	31	794	12	2	0	9	1	13	2	60
95	1	51	1200	8	1	20	9	2	17	1	94
96	2	8	1300	8	2	6	2	1	18	1	84
97	2	8	1300	8	2	6	2	1	18	2	85
98	1	15	994	8	2	8	1	1	27	1	65
99	1	15	994	8	1	7	5	1	29	1	48
100	1	15	994	10	1	6	2	1	23	1	77
101	1	15	994	9	1	4	1	1	14	1	72
102	1	15	994	7	1	7	9	1	23	1	75
103	1	15	994	8	2	10	9	1	30	2	62
104	1	15	994	8	1	3	5	1	12	1	56
105	1	15	994	9	2	0	4	1	12	2	48
106	1	15	994	8	2	1	9	1	30	1	54
107	1	15	994	7	1	2	2	1	27	1	65
108	2	22	853	11	2	11	2	2	12	1	87
109	2	22	853	12	2	23	9	1	4	1	90
110	2	22	853	12	2	23	9	1	4	1	90
111	2	22	853	12	1	10	9	3	7	2	100
112	2	22	853	7	2	0	3	2	17	2	56
113	2	22	853	.	2	33	1	1	9	2	91
114	2	22	853	7	2	4	1	2	30	1	90
115	2	22	853	11	2	0	2	2	8	1	77
116	2	22	853	11	1	12	2	1	20	1	79
117	2	22	853	11	2	0	2	2	8	1	77

	INX	TPP	SBP	TPC	SBC	TIME	Uk1_2
79	53	3	1	2	4	30	2
80	82	9	0	6	9	32	2
81	86	6	4	0	9	36	2
82	40	6	1	19	22	38	2
83	21	12	0	0	10	30	2
84	44	24	0	0	4	30	2
85	80	11	0	2	5	34	2
86	56	40	7	4	9	34	2
87	72	42	3	10	13	34	2
88	59	13	0	2	4	34	2
89	57	5	1	10	17	32	2
90	28	15	3	1	1	30	2
91	45	18	12	3	0	30	2
92	56	25	3	16	4	32	2
93	51	17	0	21	39	32	2
94	15	0	0	2	3	30	2
95	58	1	0	2	8	30	2
96	53	37	7	10	8	38	2
97	77	36	1	13	7	30	2
98	56	28	3	3	14	30	2
99	83	18	10	3	13	30	2
100	84	23	7	3	21	30	2
101	65	24	0	0	25	36	2
102	54	15	3	1	14	30	2
103	55	21	15	5	12	32	2
104	53	22	1	1	27	30	2
105	73	17	1	2	29	30	2
106	71	7	5	1	19	30	2
107	68	29	12	2	32	30	2
108	70	25	1	2	25	30	1
109	23	4	0	0	1	30	1
110	23	4	0	0	1	30	1
111	100	0	0	0	0	30	1
112	28	4	0	5	9	30	1
113	37	9	0	0	0	30	1
114	58	17	0	2	5	30	1
115	14	9	0	3	1	30	1
116	37	4	0	0	9	30	1
117	14	9	0	3	1	30	1

	LocTyp	FSM	Roll	YrGp	M_F	YrsTch	Subj	Adlts	ClasSz	AMPM	OnTsk
118	2	22	853	11	1	12	2	1	20	1	79
119	2	22	853	7	2	5	3	1	15	2	80
120	2	22	853	7	2	5	3	1	15	2	80
121	2	22	853	8	2	3	4	1	26	2	93
122	2	22	853	7	2	7	9	1	17	1	88
123	2	22	853	8	1	14	9	1	26	1	90
124	2	22	853	7	2	7	9	1	17	1	88
125	2	22	853	8	1	14	9	1	26	1	90
126	2	22	853	7	2	7	1	1	28	1	73
127	2	22	853	9	2	7	1	4	14	2	53
128	2	22	853	8	1	6	1	3	12	1	54
129	2	22	853	12	2	33	1	1	7	1	99
130	2	22	853	10	2	3	1	2	25	1	79
131	2	22	853	9	2	3	1	1	30	2	97
132	2	22	853	11	2	8	1	2	13	1	87
133	2	22	853	10	2	8	1	2	17	2	77
134	2	22	853	10	2	10	1	1	20	2	92
135	2	22	853	7	2	0	1	1	28	1	79
136	2	22	853	9	2	0	1	2	12	2	76
137	2	22	853	10	2	3	1	2	25	1	79
138	2	22	853	9	2	3	1	1	30	2	97
139	2	22	853	11	2	8	1	2	13	1	87
140	2	22	853	10	2	8	1	2	17	2	77
141	2	22	853	10	2	10	1	1	20	2	92
142	2	22	853	7	2	0	1	1	28	1	79
143	2	22	853	9	2	0	1	2	12	2	76
144	2	22	853	11	1	4	2	1	11	1	76
145	2	22	853	7	1	4	2	3	16	2	63
146	2	22	853	7	2	0	2	2	26	2	49
147	2	22	853	10	2	10	2	2	7	1	86
148	2	22	853	9	2	10	2	2	26	2	81
149	2	22	853	11	1	12	2	2	23	2	71
150	2	22	853	7	2	20	2	2	27	2	60
151	2	22	853	11	2	19	2	1	28	1	70
152	2	22	853	8	1	5	2	2	10	2	71
153	2	22	853	11	1	5	2	1	11	1	76
154	2	22	853	7	1	4	2	3	16	2	63
155	2	22	853	7	2	0	2	2	26	2	49
156	2	22	853	10	2	10	2	2	7	1	86

	INX	TPP	SBP	TPC	SBC	TIME	Uk1_2
118	37	4	0	0	9	30	1
119	28	18	0	5	12	30	1
120	28	18	0	5	12	30	1
121	17	26	17	2	12	30	1
122	13	18	1	2	6	30	1
123	13	5	0	0	13	30	1
124	13	18	1	2	6	30	1
125	13	5	0	0	13	30	1
126	12	3	0	0	7	30	1
127	18	6	0	3	20	30	1
128	24	12	1	1	15	30	1
129	43	10	0	0	0	30	1
130	15	11	0	1	20	30	1
131	30	9	0	0	6	30	1
132	60	2	0	1	2	30	1
133	20	1	0	0	2	30	1
134	15	3	0	0	0	30	1
135	29	8	3	2	16	30	1
136	47	15	0	0	2	30	1
137	15	11	0	1	20	30	1
138	21	9	0	0	6	30	1
139	60	2	0	1	2	30	1
140	20	1	0	0	2	30	1
141	15	3	0	0	0	30	1
142	29	8	3	2	16	30	1
143	47	15	0	0	2	30	1
144	19	1	0	2	8	30	1
145	25	1	0	2	12	30	1
146	40	21	1	1	18	30	1
147	31	9	0	2	11	30	1
148	23	7	0	0	4	30	1
149	35	4	0	0	4	30	1
150	27	1	0	2	22	30	1
151	38	15	1	0	2	30	1
152	26	0	0	0	2	30	1
153	19	1	0	2	8	30	1
154	25	1	0	2	12	30	1
155	40	21	1	1	18	30	1
156	31	9	0	2	11	30	1

	LocTyp	FSM	Roll	YrGp	M_F	YrsTch	Subj	Adlts	ClasSz	AMPM	OnTsk
157	2	22	853	9	2	10	2	2	26	2	81
158	2	22	853	11	1	12	2	2	23	2	71
159	2	22	853	7	2	20	2	2	27	2	60
160	2	22	853	11	2	19	2	1	28	1	70
161	2	22	853	8	1	9	2	2	10	2	71
162	2	22	853	7	2	0	3	2	25	1	52
163	2	22	853	11	2	5	3	2	13	1	79
164	2	22	853	9	1	7	3	1	27	1	91
165	2	22	853	11	1	7	3	1	27	2	46
166	2	22	853	11	1	11	3	1	30	1	56
167	2	22	853	12	1	11	3	1	3	2	91
168	2	22	853	8	2	0	3	4	27	1	81
169	2	22	853	11	2	5	3	2	13	1	79
170	2	22	853	9	1	7	3	1	27	1	91
171	2	22	853	11	1	7	3	1	27	2	46
172	2	22	853	11	1	11	3	1	30	1	56
173	2	22	853	12	1	11	3	1	3	2	91
174	2	22	853	8	2	0	3	4	27	1	81
175	2	22	853	8	2	6	4	2	18	1	77
176	2	22	853	12	2	6	4	1	12	2	90
177	2	22	853	8	2	9	4	1	27	2	72
178	2	22	853	9	2	34	4	1	28	2	77
179	2	22	853	9	1	3	4	1	25	1	58
180	2	22	853	8	1	12	4	1	13	1	88
181	2	22	853	8	2	9	4	1	27	2	72
182	2	22	853	9	2	34	4	1	28	2	77
183	2	22	853	9	1	3	4	1	25	1	58
184	2	22	853	8	1	17	4	1	13	1	88
185	2	22	853	8	2	9	5	2	27	1	86
186	2	22	853	7	2	2	5	1	17	1	35
187	2	22	853	7	2	2	5	1	16	2	54
188	2	22	853	11	2	11	5	1	8	1	92
189	2	22	853	11	2	11	5	1	13	2	84
190	2	22	853	8	2	9	5	2	27	1	86
191	2	22	853	7	2	2	5	1	17	1	35
192	2	22	853	7	2	2	5	1	16	2	54
193	2	22	853	11	2	11	5	1	8	1	92
194	2	22	853	11	2	11	5	1	13	2	84
195	2	22	853	12	2	3	9	1	12	1	89

	INX	TPP	SBP	TPC	SBC	TIME	Uk1_2
157	23	7	0	0	4	30	1
158	35	4	0	0	4	30	1
159	26	1	0	2	22	30	1
160	36	15	1	0	2	30	1
161	26	0	0	0	2	32	1
162	40	6	0	1	23	30	1
163	21	2	0	1	4	30	1
164	48	1	1	0	12	30	1
165	60	0	0	0	4	30	1
166	17	2	0	0	1	30	1
167	55	0	0	0	0	30	1
168	29	3	0	3	4	32	1
169	21	2	0	1	4	30	1
170	48	1	1	0	12	30	1
171	61	0	0	0	4	30	1
172	17	2	0	0	1	30	1
173	55	0	0	0	0	30	1
174	29	3	0	3	4	30	1
175	20	19	1	0	9	30	1
176	32	14	0	0	2	30	1
177	8	6	0	0	4	30	1
178	20	10	0	0	23	30	1
179	45	24	0	0	15	30	1
180	48	15	0	2	2	30	1
181	8	6	0	0	4	30	1
182	20	10	0	0	23	30	1
183	45	24	0	0	15	30	1
184	48	15	0	2	2	30	1
185	16	5	0	3	4	30	1
186	35	0	0	0	10	30	1
187	34	4	0	0	9	30	1
188	35	8	0	0	3	30	1
189	49	1	1	0	6	30	1
190	16	5	0	3	4	30	1
191	35	0	0	0	10	30	1
192	34	4	0	0	9	30	1
193	35	8	0	0	3	30	1
194	49	1	1	0	6	30	1
195	33	6	6	0	0	30	1

	LocTyp	FSM	Roll	YrGp	M_F	YrsTch	Subj	Adlts	ClasSz	AMPM	OnTsk
196	2	22	853	10	2	7	9	2	12	1	75
197	2	22	853	12	2	6	9	2	20	1	94
198	2	22	853	12	2	6	9	2	12	2	94
199	2	22	853	11	1	3	9	2	5	2	83
200	2	22	853	11	1	28	9	3	9	1	82
201	2	22	853	12	1	28	9	1	18	2	90
202	2	22	853	11	2	34	9	2	8	1	91
203	2	22	853	11	2	10	9	1	10	1	80
204	2	22	853	9	2	10	9	1	19	1	65
205	2	22	853	9	2	2	9	1	14	1	64
206	2	22	853	9	1	12	9	1	15	1	91
207	2	22	853	12	1	12	9	1	17	2	73
208	2	22	853	8	2	7	9	1	15	2	93
209	2	22	853	8	1	10	9	2	15	1	95
210	2	22	853	10	1	10	9	2	10	2	84
211	2	22	853	12	1	14	9	1	5	2	92
212	2	22	853	7	1	30	9	1	22	1	66
213	2	22	853	10	1	30	9	2	21	2	71
214	2	22	853	12	1	10	9	1	16	1	76
215	2	22	853	11	1	10	9	1	5	2	82
216	2	22	853	10	2	15	9	1	13	2	95
217	2	22	853	7	1	0	9	2	26	2	49
218	2	22	853	11	1	3	9	2	5	2	83
219	2	22	853	11	1	28	9	3	9	1	82
220	2	22	853	12	1	28	9	1	18	2	90
221	2	22	853	11	2	34	9	2	8	1	91
222	2	22	853	11	2	10	9	1	10	1	80
223	2	22	853	9	2	10	9	1	19	1	65
224	2	22	853	9	2	2	9	1	14	1	64
225	2	22	853	9	1	12	9	1	15	1	91
226	2	22	853	12	1	12	9	1	17	2	73
227	2	22	853	8	2	7	9	1	15	2	93
228	2	22	853	8	1	10	9	2	15	1	95
229	2	22	853	10	1	10	9	2	10	2	84
230	2	22	853	12	1	14	9	1	5	2	92
231	2	22	853	7	1	30	9	1	22	1	66
232	2	22	853	10	1	30	9	2	21	2	71
233	2	22	853	12	1	10	9	1	16	1	76
234	2	22	853	11	1	10	9	1	5	2	82

	INX	TPP	SBP	TPC	SBC	TIME	Uk1_2
196	32	18	0	1	11	30	1
197	15	10	0	0	1	30	1
198	24	4	3	0	1	30	1
199	17	15	0	1	3	30	1
200	12	1	0	0	0	30	1
201	23	7	1	3	0	30	1
202	49	16	4	0	2	30	1
203	12	2	0	0	0	30	1
204	41	5	2	1	16	30	1
205	20	3	0	0	3	30	1
206	53	9	0	0	5	30	1
207	34	1	0	1	4	30	1
208	44	4	0	0	2	30	1
209	35	9	1	3	2	30	1
210	16	0	0	2	2	30	1
211	31	5	0	0	1	30	1
212	32	6	0	2	11	30	1
213	44	6	0	0	18	30	1
214	72	3	0	0	0	30	1
215	36	2	1	1	1	30	1
216	40	6	0	0	0	30	1
217	24	5	0	1	2	30	1
218	17	15	0	1	3	30	1
219	12	1	0	0	0	30	1
220	23	7	1	3	0	30	1
221	49	16	4	0	2	30	1
222	12	2	0	0	0	30	1
223	41	5	2	1	16	30	1
224	19	3	0	0	3	30	1
225	53	9	0	0	5	30	1
226	34	1	0	1	4	30	1
227	44	4	0	0	2	30	1
228	35	9	1	3	2	30	1
229	16	0	0	2	2	30	1
230	31	5	0	0	1	30	1
231	32	6	0	2	11	30	1
232	44	6	0	0	18	30	1
233	72	3	0	0	0	30	1
234	36	2	1	1	1	30	1

	LocTyp	FSM	Roll	YrGp	M_F	YrsTch	Subj	Adlts	ClasSz	AMPM	OnTsk
235	2	22	853	10	2	23	9	1	13	2	95
236	2	22	853	7	1	0	9	2	26	2	49

	INX	TPP	SBP	TPC	SBC	TIME	Uk1_2
235	40	6	0	0	0	30	1
236	24	5	0	1	2	30	1

UK27 Dataset – parametric correlations using Pearson's *r*

		LocTp	FSM	Roll	YrGp	M_F	YrsTc	Subj	Adlts	ClaSz	AmPm	OnTsk	INX	TPP	SBP	TPC	SBC
LocTyp	Pearson Correlation	1	-.393**	.601**	-.032	.186	.146	-.081	.064	.360**	.097	.024	.199*	.159	.073	.276**	.149
	Sig. (2-tailed)		.000	.000	.743	.055	.135	.405	.515	.000	.322	.805	.040	.102	.457	.004	.125
	N	107	105	107	107	107	107	107	107	107	107	107	107	107	106	107	107
FSM	Pearson Correlation	-.393**	1	-.247*	.029	.142	.053	.298**	.217*	-.363**	-.034	.245*	.101	-.045	-.214*	-.235*	-.171
	Sig. (2-tailed)	.000		.011	.767	.148	.594	.002	.026	.000	.733	.012	.304	.646	.029	.016	.081
	N	105	105	105	105	105	105	105	105	105	105	105	105	105	104	105	105
Roll	Pearson Correlation	.601**	-.247*	1	-.129	.208*	.203*	.059	.086	.323**	.045	.023	.380**	.160	-.017	.173	.016
	Sig. (2-tailed)	.000	.011		.187	.031	.036	.544	.376	.001	.645	.811	.000	.100	.865	.075	.873
	N	107	105	107	107	107	107	107	107	107	107	107	107	107	106	107	107
YrGp	Pearson Correlation	-.032	.029	-.129	1	.124	.053	-.044	-.167	-.226*	.236*	-.161	-.155	-.340**	-.086	.079	-.081
	Sig. (2-tailed)	.743	.767	.187		.203	.584	.654	.085	.019	.015	.098	.111	.000	.381	.420	.405
	N	107	105	107	107	107	107	107	107	107	107	107	107	107	106	107	107
M_F	Pearson Correlation	.186	.142	.208*	.124	1	-.016	.218*	.004	.092	.129	-.168	.059	.082	.063	-.004	-.115
	Sig. (2-tailed)	.055	.148	.031	.203		.871	.024	.968	.347	.186	.084	.543	.401	.522	.966	.240
	N	107	105	107	107	107	107	107	107	107	107	107	107	107	106	107	107
YrsTch	Pearson Correlation	.146	.053	.203*	.053	-.016	1	-.007	.397**	-.030	.076	.358**	.262**	.092	-.046	.123	-.124
	Sig. (2-tailed)	.135	.594	.036	.584	.871		.944	.000	.761	.438	.000	.006	.343	.640	.208	.202
	N	107	105	107	107	107	107	107	107	107	107	107	107	107	106	107	107
Subj	Pearson Correlation	-.081	.298**	.059	-.044	.218*	-.007	1	.009	-.007	.005	-.052	-.137	-.151	.101	-.112	.025
	Sig. (2-tailed)	.405	.002	.544	.654	.024	.944		.929	.941	.957	.598	.159	.121	.305	.253	.802
	N	107	105	107	107	107	107	107	107	107	107	107	107	107	106	107	107
Adlts	Pearson Correlation	.064	.217*	.086	-.167	.004	.397**	.009	1	-.122	.042	.117	.045	.325**	-.084	.149	.103
	Sig. (2-tailed)	.515	.026	.376	.085	.968	.000	.929		.210	.668	.232	.645	.001	.389	.125	.291
	N	107	105	107	107	107	107	107	107	107	107	107	107	107	106	107	107
ClasSz	Pearson Correlation	.360**	-.363**	.323**	-.226*	.092	-.030	-.007	-.122	1	-.151	-.069	.084	.099	.316**	-.062	.087
	Sig. (2-tailed)	.000	.000	.001	.019	.347	.761	.941	.210		.120	.477	.389	.308	.001	.527	.371
	N	107	105	107	107	107	107	107	107	107	107	107	107	107	106	107	107
AMPM	Pearson Correlation	.097	-.034	.045	.236*	.129	.076	.005	.042	-.151	1	-.082	-.004	-.155	-.073	-.023	-.078
	Sig. (2-tailed)	.322	.733	.645	.015	.186	.438	.957	.668	.120		.403	.969	.112	.458	.813	.426
	N	107	105	107	107	107	107	107	107	107	107	107	107	107	106	107	107
OnTsk	Pearson Correlation	.024	.245*	.023	-.161	-.168	.358**	-.052	.117	-.069	-.082	1	.037	.078	-.105	-.019	-.374**
	Sig. (2-tailed)	.805	.012	.811	.098	.084	.000	.598	.232	.477	.403		.702	.425	.283	.844	.000
	N	107	105	107	107	107	107	107	107	107	107	107	107	107	106	107	107
INX	Pearson Correlation	.199*	.101	.380**	-.155	.059	.262**	-.137	.045	.084	-.004	.037	1	.196*	-.096	.116	.126
	Sig. (2-tailed)	.040	.304	.000	.111	.543	.006	.159	.645	.389	.969	.702		.043	.327	.235	.195
	N	107	105	107	107	107	107	107	107	107	107	107	107	107	106	107	107
TPP	Pearson Correlation	.159	-.045	.160	-.340**	.082	.092	-.151	.325**	.099	-.155	.078	.196*	1	.272**	.176	.201*
	Sig. (2-tailed)	.102	.646	.100	.000	.401	.343	.121	.001	.308	.112	.425	.043		.005	.070	.037
	N	107	105	107	107	107	107	107	107	107	107	107	107	107	106	107	107
SBP	Pearson Correlation	.073	-.214*	-.017	-.086	.063	-.046	.101	-.084	.316**	-.073	-.105	-.096	.272**	1	.001	.038
	Sig. (2-tailed)	.457	.029	.865	.381	.522	.640	.305	.389	.001	.458	.283	.327	.005		.996	.696
	N	106	104	106	106	106	106	106	106	106	106	106	106	106	106	106	106
TPC	Pearson Correlation	.276**	-.235*	.173	.079	-.004	.123	-.112	.149	-.062	-.023	-.019	.116	.176	.001	1	.373**
	Sig. (2-tailed)	.004	.016	.075	.420	.966	.208	.253	.125	.527	.813	.844	.235	.070	.996		.000
	N	107	105	107	107	107	107	107	107	107	107	107	107	107	106	107	107
SBC	Pearson Correlation	.149	-.171	.016	-.081	-.115	-.124	.025	.103	.087	-.078	-.374**	.126	.201*	.038	.373**	1
	Sig. (2-tailed)	.125	.081	.873	.405	.240	.202	.802	.291	.371	.426	.000	.195	.037	.696	.000	
	N	107	105	107	107	107	107	107	107	107	107	107	107	107	106	107	107

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

UK27 Dataset - non-parametric correlations using Spearman's *rho*

		LocTp	FSM	Roll	YrGp	M_F	YrsTc	Subj	Adlts	ClaSz	AmPm	OnTsk	INX	TPP	SBP	TPC	SBC
LocTyp	Correlation Coefficient	1.000	-.680**	.601**	-.018	.186	.102	-.132	-.011	.359**	.097	-.016	.239*	.139	.147	.320**	.116
	Sig. (2-tailed)	.	.000	.000	.852	.055	.297	.175	.909	.000	.322	.868	.013	.153	.133	.001	.234
	N	107	105	107	107	107	107	107	107	107	107	107	107	107	106	107	107
FSM	Correlation Coefficient	-.680**	1.000	-.491**	.051	.034	.079	.232*	.223*	-.450**	-.054	.198*	-.149	-.185	-.253**	-.321**	-.190
	Sig. (2-tailed)	.000	.	.000	.606	.729	.421	.017	.022	.000	.585	.043	.129	.058	.009	.001	.052
	N	105	105	105	105	105	105	105	105	105	105	105	105	105	104	105	105
Roll	Correlation Coefficient	.601**	-.491**	1.000	-.111	.216*	.103	-.007	.041	.342**	.038	-.020	.410**	.158	.064	.210*	-.016
	Sig. (2-tailed)	.000	.000	.	.255	.026	.292	.946	.676	.000	.701	.838	.000	.105	.516	.030	.869
	N	107	105	107	107	107	107	107	107	107	107	107	107	107	106	107	107
YrGp	Correlation Coefficient	-.018	.051	-.111	1.000	.144	.134	-.097	-.220*	-.186	.220*	-.120	-.107	-.321**	-.092	.056	-.029
	Sig. (2-tailed)	.852	.606	.255	.	.139	.169	.318	.023	.055	.023	.217	.274	.001	.349	.570	.769
	N	107	105	107	107	107	107	107	107	107	107	107	107	107	106	107	107
M_F	Correlation Coefficient	.186	.034	.216*	.144	1.000	-.030	.022	-.048	.104	.129	-.135	.058	.071	.046	.020	-.094
	Sig. (2-tailed)	.055	.729	.026	.139	.	.758	.824	.623	.285	.186	.164	.552	.468	.637	.842	.335
	N	107	105	107	107	107	107	107	107	107	107	107	107	107	106	107	107
YrsTch	Correlation Coefficient	.102	.079	.103	.134	-.030	1.000	-.084	.182	.029	.066	.425**	.236*	.021	.006	.022	-.222*
	Sig. (2-tailed)	.297	.421	.292	.169	.758	.	.390	.061	.764	.500	.000	.014	.833	.954	.820	.022
	N	107	105	107	107	107	107	107	107	107	107	107	107	107	106	107	107
Subj	Correlation Coefficient	-.132	.232*	-.007	-.097	.022	-.084	1.000	.002	-.049	.001	.000	-.150	-.193*	.093	-.056	.088
	Sig. (2-tailed)	.175	.017	.946	.318	.824	.390	.	.986	.618	.995	1.000	.123	.047	.342	.566	.365
	N	107	105	107	107	107	107	107	107	107	107	107	107	107	106	107	107
Adlts	Correlation Coefficient	-.011	.223*	.041	-.220*	-.048	.182	.002	1.000	-.150	.002	.222*	.033	.221*	-.139	-.077	-.048
	Sig. (2-tailed)	.909	.022	.676	.023	.623	.061	.986	.	.124	.986	.021	.737	.022	.156	.429	.624
	N	107	105	107	107	107	107	107	107	107	107	107	107	107	106	107	107
ClasSz	Correlation Coefficient	.359**	-.450**	.342**	-.186	.104	.029	-.049	-.150	1.000	-.139	-.094	.082	.112	.256**	.076	.085
	Sig. (2-tailed)	.000	.000	.000	.055	.285	.764	.618	.124	.	.154	.334	.402	.249	.008	.436	.382
	N	107	105	107	107	107	107	107	107	107	107	107	107	107	106	107	107
AMPM	Correlation Coefficient	.097	-.054	.038	.220*	.129	.066	.001	.002	-.139	1.000	-.051	.013	-.187	-.006	-.081	-.092
	Sig. (2-tailed)	.322	.585	.701	.023	.186	.500	.995	.986	.154	.	.604	.894	.054	.947	.409	.344
	N	107	105	107	107	107	107	107	107	107	107	107	107	107	106	107	107
OnTsk	Correlation Coefficient	-.016	.198*	-.020	-.120	-.135	.425**	.000	.222*		-.051	1.000	-.010	.009	-.128	-.070	-.397**
	Sig. (2-tailed)	.868	.043	.838	.217	.164	.000	1.000	.021	.334	.604	.	.922	.931	.190	.474	.000
	N	107	105	107	107	107	107	107	107	107	107	107	107	107	106	107	107
INX	Correlation Coefficient	.239*	-.149	.410**	-.107	.058	.236*	-.150	.033	.082	.013	-.010	1.000	.195*	-.046	.147	.141
	Sig. (2-tailed)	.013	.129	.000	.274	.552	.014	.123	.737	.402	.894	.922	.	.044	.641	.132	.147
	N	107	105	107	107	107	107	107	107	107	107	107	107	107	106	107	107
TPP	Correlation Coefficient	.139	-.185	.158	-.321**	.071	.021	-.193*	.221*	.112	-.187	.009	.195*	1.000	.352**	.226*	.215*
	Sig. (2-tailed)	.153	.058	.105	.001	.468	.833	.047	.022	.249	.054	.931	.044	.	.000	.019	.026
	N	107	105	107	107	107	107	107	107	107	107	107	107	107	106	107	107
SBP	Correlation Coefficient	.147	-.253**	.064	-.092	.046	.006	.093	-.139	.256**	-.006	-.128	-.046	.352**	1.000	.214*	.098
	Sig. (2-tailed)	.133	.009	.516	.349	.637	.954	.342	.156	.008	.947	.190	.641	.000	.	.028	.319
	N	106	104	106	106	106	106	106	106	106	106	106	106	106	106	106	106
TPC	Correlation Coefficient	.320**	-.321**	.210*	.056	.020	.022	-.056	-.077	.076	-.081	-.070	.147	.226*	.214*	1.000	.362**
	Sig. (2-tailed)	.001	.001	.030	.570	.842	.820	.566	.429	.436	.409	.474	.132	.019	.028	.	.000
	N	107	105	107	107	107	107	107	107	107	107	107	107	107	106	107	107
SBC	Correlation Coefficient	.116	-.190	-.016	-.029	-.094	-.222*	.088	-.048	.085	-.092	-.397**	.141	.215*	.098	.362**	1.000
	Sig. (2-tailed)	.234	.052	.869	.769	.335	.022	.365	.624	.382	.344	.000	.147	.026	.319	.000	.
	N	107	105	107	107	107	107	107	107	107	107	107	107	107	106	107	107

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

UK1 Dataset – parametric correlations using Pearson's *r*

		YrGp	M_F	YrsTch	Subj	Adlts	ClasSz	AmPm	OnTsk	INX	TPP	SBP	TPC	SBC
YrGp	Pearson Correlation	1	-.186*	.301**	.193*	-.126	-.477**	-.017	.390**	.189*	-.201*	-.030	-.349**	-.472**
	Sig. (2-tailed)		.035	.001	.029	.155	.000	.850	.000	.033	.023	.738	.000	.000
	N	128	128	128	128	128	128	128	128	128	128	128	128	128
M_F	Pearson Correlation	-.186*	1	-.151	-.291**	.023	.094	-.076	.065	-.182*	.243**	.151	.050	.141
	Sig. (2-tailed)	.035		.087	.001	.796	.291	.390	.463	.039	.006	.088	.577	.111
	N	128	129	129	129	129	129	129	129	129	129	129	129	129
YrsTch	Pearson Correlation	.301**	-.151	1	.297**	-.147	-.123	.008	.270**	.069	-.048	-.024	-.182*	-.034
	Sig. (2-tailed)	.001	.087		.001	.097	.166	.932	.002	.440	.592	.785	.039	.703
	N	128	129	129	129	129	129	129	129	129	129	129	129	129
Subj	Pearson Correlation	.193*	-.291**	.297**	1	-.177*	-.320**	-.047	.189*	.023	-.091	.068	-.102	-.277**
	Sig. (2-tailed)	.029	.001	.001		.044	.000	.598	.032	.798	.306	.447	.250	.002
	N	128	129	129	129	129	129	129	129	129	129	129	129	129
Adlts	Pearson Correlation	-.126	.023	-.147	-.177*	1	-.039	.029	-.113	-.080	-.030	-.089	.303**	.109
	Sig. (2-tailed)	.155	.796	.097	.044		.665	.740	.202	.370	.735	.315	.000	.219
	N	128	129	129	129	129	129	129	129	129	129	129	129	129
ClasSz	Pearson Correlation	-.477**	.094	-.123	-.320**	-.039	1	-.021	-.345**	-.134	.099	.067	.021	.428**
	Sig. (2-tailed)	.000	.291	.166	.000	.665		.815	.000	.130	.263	.449	.815	.000
	N	128	129	129	129	129	129	129	129	129	129	129	129	129
AMPM	Pearson Correlation	-.017	-.076	.008	-.047	.029	-.021	1	-.063	.000	-.137	-.016	-.058	-.056
	Sig. (2-tailed)	.850	.390	.932	.598	.740	.815		.480	.999	.121	.855	.511	.532
	N	128	129	129	129	129	129	129	129	129	129	129	129	129
OnTsk	Pearson Correlation	.390**	.065	.270**	.189*	-.113	-.345**	-.063	1	.063	.093	.152	-.023	-.395**
	Sig. (2-tailed)	.000	.463	.002	.032	.202	.000	.480		.482	.296	.085	.798	.000
	N	128	129	129	129	129	129	129	129	129	129	129	129	129
INX	Pearson Correlation	.189*	-.182*	.069	.023	-.080	-.134	.000	.063	1	.053	-.013	-.189*	-.019
	Sig. (2-tailed)	.033	.039	.440	.798	.370	.130	.999	.482		.552	.888	.032	.828
	N	128	129	129	129	129	129	129	129	129	129	129	129	129
TPP	Pearson Correlation	-.201*	.243**	-.048	-.091	-.030	.099	-.137	.093	.053	1	.328**	.191*	.284**
	Sig. (2-tailed)	.023	.006	.592	.306	.735	.263	.121	.296	.552		.000	.030	.001
	N	128	129	129	129	129	129	129	129	129	129	129	129	129
SBP	Pearson Correlation	-.030	.151	-.024	.068	-.089	.067	-.016	.152	-.013	.328**	1	.071	.083
	Sig. (2-tailed)	.738	.088	.785	.447	.315	.449	.855	.085	.888	.000		.423	.352
	N	128	129	129	129	129	129	129	129	129	129	129	129	129
TPC	Pearson Correlation	-.349**	.050	-.182*	-.102	.303**	.021	-.058	-.023	-.189*	.191*	.071	1	.207*
	Sig. (2-tailed)	.000	.577	.039	.250	.000	.815	.511	.798	.032	.030	.423		.019
	N	128	129	129	129	129	129	129	129	129	129	129	129	129
SBC	Pearson Correlation	-.472**	.141	-.034	-.277**	.109	.428**	-.056	-.395**	-.019	.284**	.083	.207*	1
	Sig. (2-tailed)	.000	.111	.703	.002	.219	.000	.532	.000	.828	.001	.352	.019	
	N	128	129	129	129	129	129	129	129	129	129	129	129	129

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

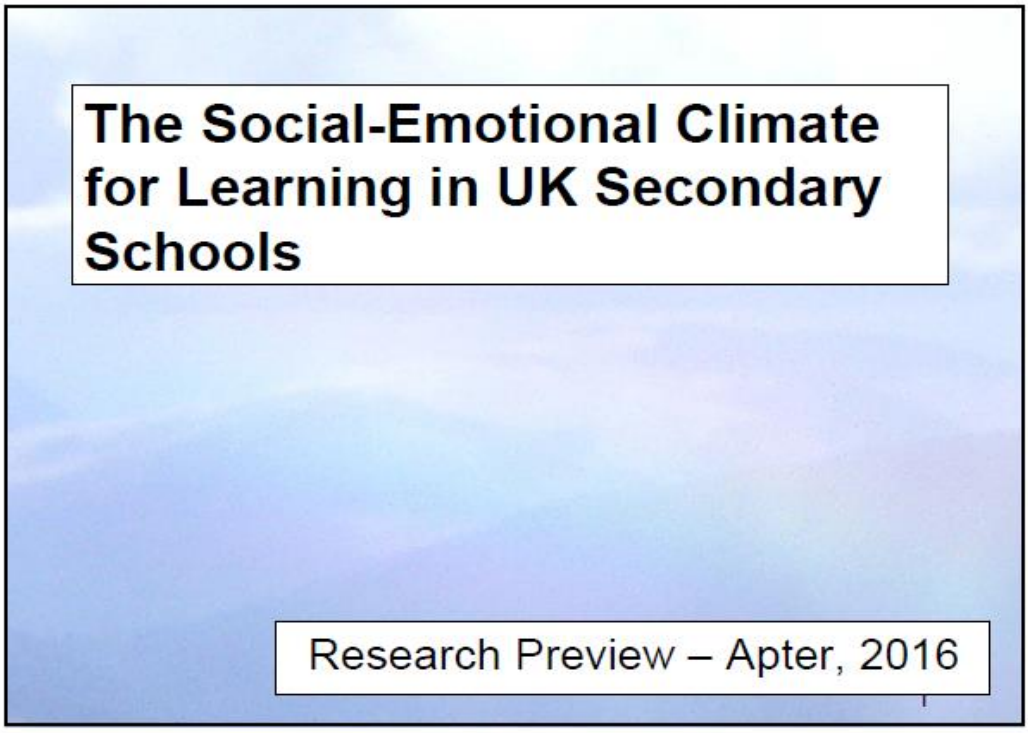
UK1 Dataset – non-parametric correlations using Spearman's *rho*

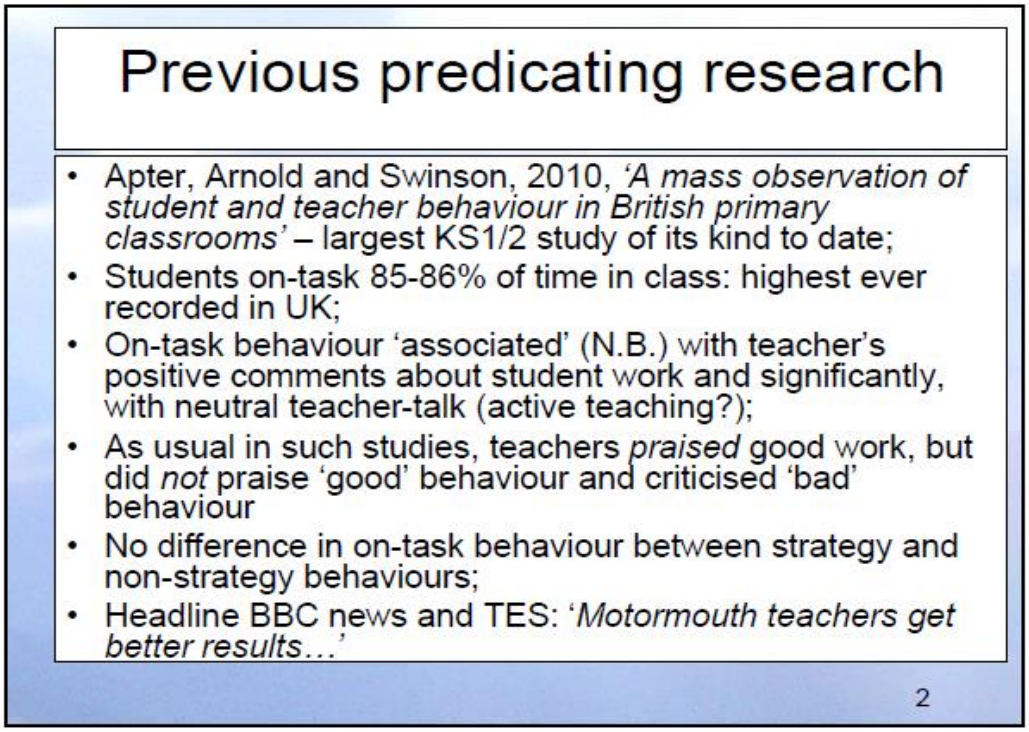
		YrGp	M_F	YrsTch	Subj	Adlts	ClasSz	AmPm	OnTsk	INX	TPP	SBP	TPC	SBC
YrGp	Correlation Coefficient	1.000	-.194*	.402**	.182*	-.124	-.475**	-.016	.338**	.129	-.192*	-.009	-.331**	-.585**
	Sig. (2-tailed)	.	.028	.000	.040	.164	.000	.854	.000	.148	.030	.923	.000	.000
	N	128	128	128	128	128	128	128	128	128	128	128	128	128
M_F	Correlation Coefficient	-.194*	1.000	-.262**	-.293**	.036	.093	-.076	.081	-.153	.296**	.111	.008	.129
	Sig. (2-tailed)	.028	.	.003	.001	.686	.294	.390	.364	.082	.001	.211	.928	.145
	N	128	129	129	129	129	129	129	129	129	129	129	129	129
YrsTch	Correlation Coefficient	.402**	-.262**	1.000	.301**	-.236**	-.142	.003	.308**	.146	-.070	.001	-.256**	-.233**
	Sig. (2-tailed)	.000	.003	.	.001	.007	.108	.970	.000	.099	.430	.987	.003	.008
	N	128	129	129	129	129	129	129	129	129	129	129	129	129
Subj	Correlation Coefficient	.182*	-.293**	.301**	1.000	-.215*	-.274**	-.049	.183*	.005	-.054	.124	-.084	-.260**
	Sig. (2-tailed)	.040	.001	.001	.	.014	.002	.581	.038	.953	.544	.162	.346	.003
	N	128	129	129	129	129	129	129	129	129	129	129	129	129
Adlts	Correlation Coefficient	-.124	.036	-.236**	-.215*	1.000	-.093	.053	-.145	-.135	.009	-.084	.334**	.093
	Sig. (2-tailed)	.164	.686	.007	.014	.	.292	.552	.101	.127	.922	.343	.000	.296
	N	128	129	129	129	129	129	129	129	129	129	129	129	129
ClasSz	Correlation Coefficient	-.475**	.093	-.142	-.274**	-.093	1.000	-.010	-.335**	-.107	.097	.055	.025	.479**
	Sig. (2-tailed)	.000	.294	.108	.002	.292	.	.908	.000	.229	.273	.539	.782	.000
	N	128	129	129	129	129	129	129	129	129	129	129	129	129
AMPM	Correlation Coefficient	-.016	-.076	.003	-.049	.053	-.010	1.000	-.039	.019	-.169	-.129	-.110	-.083
	Sig. (2-tailed)	.854	.390	.970	.581	.552	.908	.	.664	.833	.056	.144	.215	.352
	N	128	129	129	129	129	129	129	129	129	129	129	129	129
OnTsk	Correlation Coefficient	.338**	.081	.308**	.183*	-.145	-.335**	-.039	1.000	.070	.166	.148	-.090	-.429**
	Sig. (2-tailed)	.000	.364	.000	.038	.101	.000	.664	.	.433	.061	.094	.311	.000
	N	128	129	129	129	129	129	129	129	129	129	129	129	129
INX	Correlation Coefficient	.129	-.153	.146	.005	-.135	-.107	.019	.070	1.000	.028	.170	-.183*	.018
	Sig. (2-tailed)	.148	.082	.099	.953	.127	.229	.833	.433	.	.757	.054	.038	.841
	N	128	129	129	129	129	129	129	129	129	129	129	129	129
TPP	Correlation Coefficient	-.192*	.296**	-.070	-.054	.009	.097	-.169	.166	.028	1.000	.270**	.183*	.241**
	Sig. (2-tailed)	.030	.001	.430	.544	.922	.273	.056	.061	.757	.	.002	.038	.006
	N	128	129	129	129	129	129	129	129	129	129	129	129	129
SBP	Correlation Coefficient	-.009	.111	.001	.124	-.084	.055	-.129	.148	.170	.270**	1.000	.154	.099
	Sig. (2-tailed)	.923	.211	.987	.162	.343	.539	.144	.094	.054	.002	.	.082	.263
	N	128	129	129	129	129	129	129	129	129	129	129	129	129
TPC	Correlation Coefficient	-.331**	.008	-.256**	-.084	.334**	.025	-.110	-.090	-.183*	.183*	.154	1.000	.276**
	Sig. (2-tailed)	.000	.928	.003	.346	.000	.782	.215	.311	.038	.038	.082	.	.002
	N	128	129	129	129	129	129	129	129	129	129	129	129	129
SBC	Correlation Coefficient	-.585**	.129	-.233**	-.260**	.093	.479**	-.083	-.429**	.018	.241**	.099	.276**	1.000
	Sig. (2-tailed)	.000	.145	.008	.003	.296	.000	.352	.000	.841	.006	.263	.002	.
	N	128	129	129	129	129	129	129	129	129	129	129	129	129

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

DISCLAIMER These presentation slides contain uncorrected initial data (pre-peer review and examiner scrutiny) that differ from final findings in the text-body of this thesis. These data findings should not be used in any subsequent publication.





Current research

- This data was collected using *MICRO* – a Mixed-Interval Classroom Observation time-sampling tool developed for Educational Psychologists;
- Part 1: National sample (NaS) of 107 academic lesson observations in 2013 in 26 UK secondary schools / academies;
- Part 2: 129 academic lesson observations Spring and Summer terms of 2014 of each teacher, a.m. and p.m. in one *anonymous* 'bog-standard' secondary school (BoS) judged 'inadequate' (Ofsted inspection, December, 2013);
- Part 3: statistically analyse Part 1 and 2 datasets;
- Part 4: compare and contrast Part 1 and 2 data-sets and analyses;
- Part 5: share results with, and consult SLT in BoS, about meaning and utility of results.

3

List of variables considered

- City, suburban or rural;
- Free school meals;
- Roll;
- Year group;
- No. of years teaching experience;
- Teacher gender;
- Time of day;
- Subject;
- No. adults in class;
- Class size;

Student on-task time %;
Neutral teacher-talk time %;
No. positive comments about academic performance by teacher;
No. positive comments about social-behaviour
No. critical comments about academic performance
No. critical comments about social-behaviour.

4

National sample (NaS): n=107

Tallies of types of comments by
NaS teachers, comparing means per lesson:

	Teachers' positive comments…	Teachers' negative comments…	Ratio:
…about students' work:	16.85	5.4	3:1
…about students' behaviour:	2.91	12.32	1:4
Ratio:	5:1	1:2	Total Ratio (+ive : -ive) 20 :18

5

Bog Standard (BoS) sample: n=129

Tallies of types of comments by
BoS Teachers, comparing means per lesson:

	Teachers' positive comments…	Teachers' negative comments…	Ratio:
…about students' work:	6.92	0.49	14:1 (NaS 3:1)
…about students' behaviour:	0.86	6.44	1:6 (NaS 1:4)
Ratio:	7:1 (NaS 5:1)	1:13 (NaS 1:2)	Total Ratio (+ive : -ive) 8 : 7 (20:18)

6

Bog Standard (BoS) sample: n=129

Tallies of types of comments by teachers, comparing means per lesson:

	Teachers' positive comments…	Teachers' negative comments…
…about students' work:	6.92 (NaS 16.85) (KS1/2 14.62)	0.49 (NaS 5.4) (KS1/2 4.61)
…about students' behaviour:	0.86 (NaS 2.91) (KS1/2 2.85)	6.44 (NaS 12.32) (KS1/2 9.49)

7

Percentage-time neutral teacher-talk (active teaching?) in academic lessons:

NaS	52.77%
BoS	31.88%
KS1/2	28.30%

8

Despite significant differences (lower scores) in teacher verbal behaviour in all categories, BoS students are *not* significantly *less* on-task than NaS students:

NaS on-task mean =	78.09 %
BoS on-task mean =	77.66 %

Why?

9

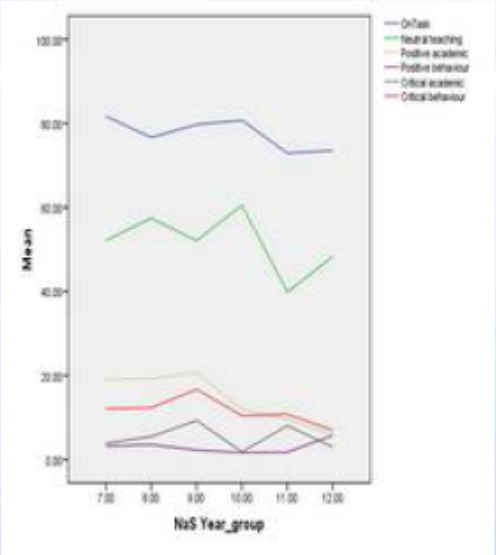
Do we envy the on-task mean enjoyed by primary schools - particularly in BoS Y7?

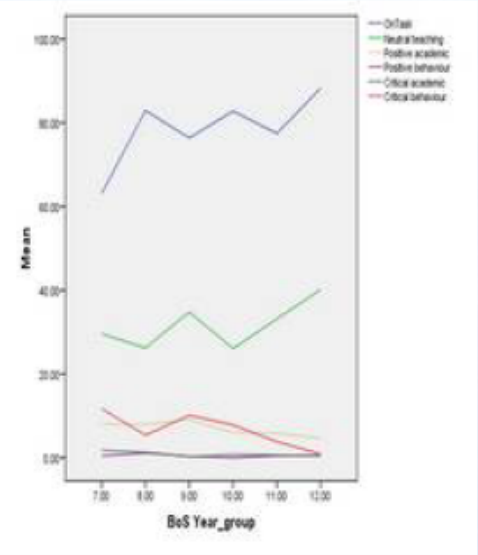
BoS on-task mean =	78.09 %
BoS Y7 on-task mean=	62.89%
KS1/2 on-task mean =	85.76 %

Why the significant differences?

10

Differences in year groups





11

Strong associations: causal and correlational (as shown by regression, correlation, parametric and non-parametric statistical tests):

- Critical comments by both NaS and BoS teachers about student social-behaviour are strongly *causally* associated with higher levels of on-task behaviour;
- Year group is *correlationally* associated with on-task behaviour; weakly with the NaS sample, but more significantly with the BoS sample.
- The number of years both NaS and BoS teachers have been teaching are *correlationally* associated with higher levels of on-task behaviour and with higher percentages of teacher-talk in lessons.

12

Two missing correlations in secondary schools *in comparison* to primary schools…

- Insignificant correlation / association between how much a teacher talks and student time on-task;
- Insignificant correlation / association between use of academic praise and student time on task;
- But remember, we are getting significantly lower on-task behaviour in both NaS and BoS samples, compared to the Primary school sample.

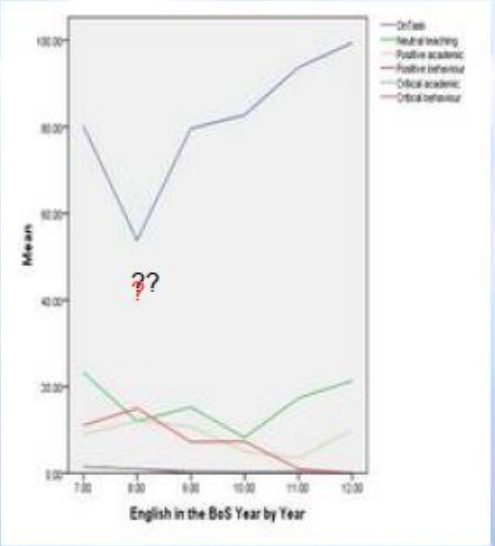
13

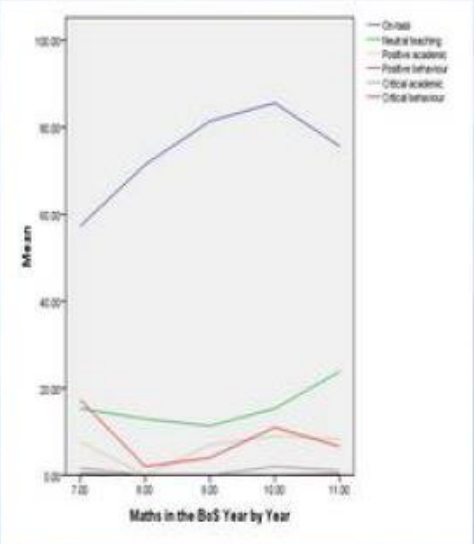
English and maths in the BoS

- Students were significantly more on-task in English (83.37%) than in Maths (71.18%).
- English lessons were characterised by higher levels of positive comments about students' work and higher levels of criticism of behaviour.
- Maths lessons were characterised by similar levels of positive comments about students' work but comparatively higher levels of criticism of behaviour.

14

English & maths in the BoS year by year





15

The social / emotional climate for learning hypothesis…

- Learning in the academic classroom is *in-part* (N.B.) a function of the social / emotional climate for learning;
- This climate can be characterised, measured and quantified by examining the way that teachers talk / feedback in the classroom and how this effects the time that students stay on-task; e.g. BoS English;
- The effect size (ES) of a teacher's use of feedback is 1.13 - equivalent to 2 grade leaps at GCSE (Prof. John Hattie, 1999);
- This hypothesis does *not* examine the curriculum; what is actually taught; the materials a teacher uses to teach with; or the teaching techniques that they use to explain and embed new knowledge; or the student knowledge / culture baseline (ES=1.04) that teachers are working with, e.g. CDPDs (culturally deprived and parentally disadvantaged).

16

Questions

1. Students are on-task in the BoS as much as students in the NaS, but it appears that they are not being taught as urgently / as verbally / as actively. How would you encourage BoS teachers in all years to teach more verbally and actively and improve the social emotional climate for learning in all classrooms?
2. How would you encourage BoS teachers to use more strategic and effective behavioural criticism / management with different year groups?
3. Are the social emotional climates of maths classrooms improving? If so, in what way and in what years? Could we find out by measuring the use of feedback and on-task percentages perhaps?
4. Could you improve on-task levels (from77% to 85%) in Y7, and Y8, by behaving more like primary school teachers (more praise for good work and good behaviour, and more verbally-vigorous teaching behaviour)?

Thank you for taking part!

18

UK1 School SLT Meeting, 9.3.16 Transcription – Iterative Learning Conversation with Foucauldian thematic discourse-analysis using three-pass coding

Pass #1 – ILCS Coding

The first coding pass through the transcribed text was used to analyse and annotate each line using Iterative Learning Conversations codes (ILCS; Apter, 2014):

Iterative Learning Conversations ILCS categorical codes		
Primary	**A1: Agents / audiences**	Lists of people involved in the change-process[a] – agentically (agents) and non-agentically (audiences – stake-holders and bystanders) in talking/communicating/behaving within the transcribed/annotated ILCS. Agents and audiences are described in respect of identities *and* in respect of their context – the historical, cultural and political "givens" that are assumed by the perceiver/recorder[b] to apply to them.
	A2: Attributions / agency	The agency (evident intentions) of agents; and the assumed knowledge expressed by both agents and audiences – including speech-*acts* (**ACT**, **Iy** or **Py**), represented usually as verbatim adjacency pairs and perceptions from the standpoint of the perceiver/recorder (usually the psychologist) of others' attributions. "Attribution" is used in the Heider (1958) *Attribution Theory* sense of contextualised meanings.
Subsidiary	**P1: Possibilities**	Possibilities of change (problems) that are special types of attributions identified by each agent, including possibilities/problems that are jointly identified by agents or audience members.
	P2: Propositions	These are tentatively proposed functional relationships – usually between possibilities of change (P1) and the elements that make up a context. Differing from the concept of a hypothesis, a *proposition* is a sentence comprised of words that attribute a causal relationship existing between different pieces of information, for example, *the boy feels sad because his mother has died*, or *if we put a plaster on the cut, it will stop bleeding*. In a given ILCS account, there might be multiple related propositions that are deemed worthy of testing.
	T1: Tests	Quasi-objective tests, checks and trials that produce ostensible evidence (E1) that can be associated with propositions (P1) in order to expose their utility. Tests might include behavioural experiments (trying a new behaviour out), data from standardised assessments including questionnaires, graded tasks or exposure to new contexts (for example, trial placements), and verbal checking, for example, *how is that so?*
	T2: Timescales	Historical strings of related events, *timed* and *dated*, including when events, contexts and behaviours have occurred, when assessments and interventions were completed, and when they are planned to occur in the future.
	E1: Evidence	Quasi-objective information and measures – including T1, observations, test and experiment results; and emergent information from trial placements and changed contexts. Evidence should include judgements by the recorder of ecological and systemic validity from *their standpoint*.
	E2: Estimates	Quasi-objective information that is not known has been estimated, projected or regressed from E1 and the identification of information that is still required. E2 is usually attributed information and thus can be expressed as a factor of A2, thus: A2(E2).
	R1: Reflections	Emotional, social and intellectual *post hoc* reflections of agents and audiences including the reflections of the ILCS recorder (usually the psychologist) which might be informed by cited research – expressed optionally as A2(R1).
	R2: Responses	A punctuation at a hypothetical endpoint of an ILCS, as it has been signified by the recorder; including planned next steps, reboots of ILCS, repeats of sections (re-iterations) and the reasons for so doing attributed by agents/audiences. A reiteration might branch back to an earlier specified point in the ILCS. A response at the end of an ILCS might include a planned review cycle or case-closure.

[a]Every socio-historical process is a *change* process, including organisational management, personal therapy, and governance.
[b]A perceiver/recorder is an ILCS practitioner who might also be a psychologist, management consultant, therapist or politician.

Pass #2 – Speech Act Coding

The second pass employed a variation of Foucauldian discourse-analysis (Foucault, 1969; Apter, 2014) using Speech Act codes for locutionary, illocutionary and perlocutionary elements (Austin, 1962).

Foucault's (1962) method emphasises the power relationships assumed and utilised in discourse. Austin (1962) proposes that verbal behaviour is mostly comprised of a series of performatives intended to have an effect on others, in contrast to passive statements. In order to annotate how power is utilised in conversations, ILCS uses Austin's Speech Act Theory to categorise performatives into three integral elements called speech acts:

(1) A locutionary act (**ACT**): the performance of an utterance: the actual utterance and its ostensible meaning, comprising the verbal, syntactic and semantic aspects of any meaningful utterance.
(2) An illocutionary act (**Iy**): the pragmatic force of the utterance, in respect of its intended significance as a socially-valid verbal action.
(3) A perlocutionary act (**Py**): its intentional effect, for example, persuading, convincing, scaring, enlightening, inspiring, or otherwise getting someone to do or realise something, whether intended or not. The perlocutionary act is implicit in the illocutionary act but is gauged once a behavioural response, usually a verbal response, has been made by a respondent.

The second-pass coding analysis was used to reveal and describe the author's perception of power-plays within the discourse. The forcefulness of a selection of speech acts is coded in bold script, e.g. **Iy** for illocutionary and **Py** for perlocutionary. These codes and descriptive annotations are placed between braces: { and }. The selection and forcefulness of lines in the transcript is necessarily made, judged and coded by the author from the author's standpoint.

Pass #3 - Thematic Coding

The third pass employed thematic colour-codes (Harding, 2013) for broad themes identified by the author as having particular relevance for the hypothesis and research questions of the current research. Thematic analysis of the transcript revealed *n* themes that were colour-coded, thus:

Method	Red	Historical / previous research	Pink
Dependent variable (On-Task)	Green	Independent: teacher verbal behaviour	Lime
UK1 Results and solutions	Blue	UK27 Results	Sky
UK1 v. UK27 Comparison	Violet	Independent variables: non-verbal	Indigo
Ofsted	Orange	Omissions and alternative (non-research) solutions	Yellow

Transcription with Three-pass Coding analysis annotations

A1: Location staff conference room with author (AUT) and 7 members of the school's Senior Leadership Team (SLT) present AUT, Head Teacher (HT) [female (f.)], Assistant Headteacher 1 (AH1) [male (m.)], AH2 [m.], Class Teacher 1 (CT1) [m.], AH3 [m.], Head of Key Stage 4 (HKS4) [f.], Head of Key Stage 3 (HKS3) [f.]

1. **A1:** AUT **A2: ACT:** *Anne* [HT] *has invited me to talk to you today about this research which has two strands… One of which relates very closely to this school… You probably know a bit about it.* [Pause. Audible, '*mm*' of agreement.] *So who's who?* {**Iy** AUT establishing leadership of discussion and power relationship with attendees apparently invitationally.} *Can we go round?*

2. **A1:** AH1 **A2: ACT:** {**Py** compliant response maintaining AUT's power-grab.} *John* [surname]*… Assistant head…*

3. **A1:** AH2 **A2: ACT:** {**Py** compliant response supporting AUT's developing meeting status.} *Chris* [surname]*… Assistant Head…*

4. **A1:** AUT **A2: ACT:** {**Iy** AUT further provocation of power relationship.} *With particular responsibilities?*

5. **A1:** AH1 **A2: ACT:** *Yeh, years eight and nine…*

6. **A1:** AH2 **A2: ACT:** *Year seven…*

7. **A1:** CT1 **A2: ACT:** *Sean* [surname]… {**Py** CT1 establishing he's bottom of pecking order.} *Just a teacher.*

8. **A1:** AUT **A2: ACT:** *Just a teacher…* [AUT + others chuckles warmly, '*awh!*', e.g. AH1, AH2, join in chuckling, and '*aah'*-ing, in mock commiseration.]

9. **A1:** AH3 **A2: ACT:** {**Py** compliance.} *Jamie [surname]… Assistant Headteacher… Pastoral.*

10. **A1:** HT **A2: ACT:** {**Py** compliance.} *Anne* [surname]*… Head Teacher.*

11. **A1:** HKS4 **A2: ACT:** {**Py** compliance.} *Bernadette* [surname]*… Head of Key stage four.*

12. **A1:** HKS3 **A2: ACT:** {**Py** compliance.} *Doris* [surname]*… Key stage three.*

13. **A1:** AUT **T1: ACT:** {**Iy** AUT self-deprecating, saying he's not a threat.} *Okay… There is something very boring about statistics… And I think… The more honest they… The statistics, are, the more boring they seem… Unfortunately… And you should never start a talk by saying how boring it's going to be…* [group laughter] *And indeed, I* [emphasis on '*I*'] *think you are going to find this* [this statistics] *quite interesting… Because they apply to, to, to… You…*

14. **P1:** [Pause, intake of breath. Continues.] *The thing that strikes me about a lot of the stuff; I've been pawing over it again over the last two days… In putting this together… For you. One thing that is evident, is that there is nothing very clear in this. And I come to you, in the hope that you can clarify some of this…* {**Iy** AUT repeating self-deprecation - no threat.} *Because… After all, you are the experts, really. I know that everyone who comes to talk in this way says that to people… But really… You really are. But now, I'm out of the game… I just do work for the British Psychological Society. I used to work in this city too. Obviously, for a number of years… But, um… I'm now a little bit out of touch with schools in the city* [emphasising modesty of position and superior knowledge of audience]. *I'm really perplexed by some of the data that's popped out here.*

15. **E1/R1:** [Pause, intake of breath. Continues.] {**Iy** AUT no longer self-deprecating - AUT saying he's an authority in these matters.} *In two thousand and eight, I did a big piece of research around primary schools in the UK… And I got… We… I put together a little team… My two friends and I got known for… Publishing the largest study, of primary… The largest observational study of primary classrooms. And the date of publishing was two thousand and ten. And I discovered that kids in primary classrooms are really on task… Really on task, for a very high degree of the time that they are in lessons… Which is this figure…* [points to PowerPoint Slide 2 (PS2)] *For eighty-five point six percent of the time. Which was… Probably… Ten percent higher than had been previously recorded… Going back to different studies… So, so my thought was… Perhaps… That maybe the National Curriculum has made the big difference… I wasn't sure… So the big headlines from that research were… That… Um… On task behaviour was associated with teachers' positive comments about students' work…* {**Iy** AUT developing his authority in these matters.} *There were no surprises there. And that, significantly, neutral teacher-talk. Was associated with… This is using correlational statistics. So you can't say 'caused-by'… You must say 'associated-with' correlationally. But some of the statistics I have used… Er… When doing the present study… Are more… Use analysis of regression, where you can say… With more confidence … That there is a, sort of, causal relationship between two variables. So we will get on to that…*

16. **R1:** [Pause, intake of breath. Continues.] *So as usual…* {**Iy** AUT developing his authority in these matters.} *There have been so many studies… That go right back to the nineteen twenties… Teachers praise good work… But do not praise good behaviour. That's the banner headline that has always been there… And in the seventies and eighties, there was a big move to try and change that. To try and get teachers to praise good behaviour… Social behaviour more... Rather than just praising good work. Teachers tend to criticise bad behaviour… So there was no surprises there… But there was a surprise that there was no difference between on-task behaviour between strategy* [National Curriculum literacy and numeracy lessons] *and non-strategy lessons…* {**Iy** AUT emphasising his authority in these matters.} *I got the headlines on the BBC and in the Times Ed., and this is the headline…* [PS2] *'Motormouth teachers get better results.' Teachers who were talking a lot… And working very energetically in the classroom… Were getting better 'results' in terms of kids in class being more 'on-task'. Then, everybody was happy to convert this to results, meaning 'scores on the door'* [published school league tables]. *Obviously, this wasn't always the case.*

17. **T1:** [Pause, intake of breath. Changes overhead display to PS3]. *So the current research is a bit more complex…* {**Iy** AUT continuing to emphasise authority in these matters.} *It's to do also with my doctorate… I am in the last stage of writing this up. So it's quite a big study again. So I did a large UK survey… So I got about one hundred and seven academic lesson observations… Of secondaries… Of… Of classes in secondary schools… In twenty seven secondary schools… Mainly in the larger conurbations… Um… London, Birmingham, Liverpool, Glasgow, Cardiff and Manchester… Um… But also in many other places… But then… I used this particular time-sampling technique… One of my bugbears… Going back to Rutter's nineteen seventy nine study, Fifteen Thousand Hours…* {**Iy** AUT continuing to emphasise authority in these matters but with tentative tone.} *Um, do you know that study?*

18. **A1:** ALL **A2: ACT:** [General murmur from group.] {**Py** Compliant group response.} *No, no, not yet…* [Etc.]

19. **A1:** AUT **E1/R1: ACT:** [Pause, intake of breath. Continues.] *It was a landmark study of seven secondary schools… And it probably… If you read it carefully… And I have to be a bit careful here… Out of that study, Ofsted was invented.*

20. **A1:** HT **A2: ACT:** *No… Not sure that I knew that…*

21. **A1:** AUT **P2: ACT:** [Continues.] *A lot of the things that Ofsted have beaten their drum about over the years have come out of that study…*

22. **R1:** [Continues.] *Unfortunately many commentators wrote that the study was statistically flawed… This didn't stop the study being incredibly… Incredibly influential… And on both sides of the Atlantic… It was actually imported into North America, where it was also used as a very influential… Piece of evidence… To say…*

23. **T2/R1:** [Continues with pause (for dramatic effect?)] *And here's the bit that you all know… To say… That poverty and cultural impoverishment are not an obstacle to the improvement of a school's academic results… And no excuse for the failure of a school to continually improve those results… That was in Rutter… He looked at seven secondary schools in these inner London deprived communities… Some were working well… And some were not working so well… And he said that the results were primarily dependent on leadership… And then after that, upon every other within-school factor that could be made the responsibility of that leadership…*

24. **R1:** [Continues with pause (for dramatic effect?)] *Sounds like Ofsted's checklist, doesn't it? No account taken… Of the catchment… Or the baseline that a school is working with when they take students in… In year seven…*

25. **E1/R2:** [Continues.] {**Iy** AUT continuing to emphasise authority in these matters.} *But many statisticians came out of the woodwork and said… But this is rubbish… Statistical conclusions by Rutter… Um… Should not have been made. The way it was done... Only seven schools... Lots of idiosyncratic idiographics… And variables… Not taken in to account… And all sorts of poor statistical techniques… Used to make unsustainable inferences… But anyway…*

26. **P1/P2:** [Continues.] {**Iy** AUT Hyperbole joke – emphasising authority by implied self-depracation –my study will NOT be perfect but it will be authentic.} *So of course… My study… Our study… Is…* [Pause (for dramatic effect?)] *Perfect!*

27. **A1:** ALL **A2: ACT:** {**Py** Group-mirth. Compliance.} [General laughter from group.] *Of course.*

28. **A1:** AUT **A2/T1: ACT:** [Continues.] {**Iy** AUT continuing to emphasise authority in these matters.} *So part two… So I did part one, and organised observations of one hundred and seven lessons up and down the UK… And then Anne* [HT] *said… Very kindly… And said that I could observe… Every teacher… In a lesson in the morning and in a lesson in the afternoon. Which was very kind of her… And you. And.. Um… That really got started in the early part of two thousand and fourteen. This process was assisted… Handily… By the fact that I had a young psychologist help me… He was doing his own PhD… And wanted to observe some UK classrooms… And he was willing to transplant himself in to this city. And his name is now Doctor* [surname]. *So well done him… And you all probably met him and knew him when he was here…*

29. **A1:** ALL **A2: ACT:** [General murmurs of agreement and pleasure.]

30. **A1:** AUT **A2/R1: ACT:** [Continues.] *Yes, it's good… He got his doctorate out of what he did here… He used some of the data he collected here for his own purpose… Which was fine and we agreed for that to happen… Anne and myself. With our agreement… But he's not publishing anything without a further agreement with myself and Anne. So I still have control over that… And the data is of course presented anonymously in his doctorate…*

31. **E1:** [Continues.] {**Iy** AUT developing authority by proxy – association with Ph.D student's work.} *So now, with Dr* [surname]*'s assistance I have an impressive dataset from your school. It is huge. I have one hundred and twenty nine academic lesson observations… This is leaving out P.E., music and I.T. I have observations of those subjects… But I have left them out… For now… But I still have the one hundred and twenty nine academic lesson*

observations using this time-sampling protocol. {**Iy** AUT large 'n' used to underpin authority.} *It is still a huge dataset… And it has a direct relevance for your school. My intention was to compare the two datasets and do a comprehensive statistical analyses on it… And then come to you to ask… Armed with this initial presentation that I have brought with me… Is this comparison in any way useful to you? And if so, what use would you make of it? ... And are there things in the future that we can do together, that might be useful to your school? …*

32. **A1:** AUT **A2: ACT:** [Uses clicker device to move to PS4.] {**Iy** AUT using humour to enrol SLT audience.} *In this presentation I have referred to you 'Bog Standard Standard School… Which I hope is ironic enough not to offend or insult any of you?*

33. **A1:** ALL **A2: ACT:** [Murmurs, quiet laughter.]

34. **A1:** AUT **T2: ACT:** [Continues.] *And… You were judged inadequate in your Ofsted inspection, end of two thousand and thirteen… That coincided with this research and the data-collection from your school. But strangely, plays into one of the more implicit themes of my research… Which is to question how classroom observations… Like Ofsted use… Are done. And hopefully ask some challenging questions…* [Pause.]

35. **P1:** [Continues.] {**Iy** Use of irony to empathise with, and continue to enrol SLT.} *Sometimes, I get this strange notion that Ofsted are political… Strangely… More so, than rigorous… But obviously, I am probably wrong about this…* [Pause.]

36. **A1:** ALL **A2: ACT:** {**Py** Compliance.} [Much laughter of agreement.]

37. **A1:** AUT **T1: ACT:** [Continues.] {**ly** AUT next 3 lines: 37,38 and 39 are about giving information that will be of interest and engage this audience because of specific relevance.} *Okay… So… So… Moving on…* [Clicks to PS4] *So these are the variables… That I considered… So I didn't have any 'rural secondary schools'… I hadn't thought of that… There aren't many that are built in the countryside… Most are in city or suburban settings. Or largish towns… So…* [Reading from PS4.] *Free school meals; roll, year group… No of years teaching experience… That was a new one that I hadn't considered before… In the previous research… It was interesting to look here* [in this school] *and in the UK sample, at each teacher that was observed… And ask, 'how long have you been teaching?' It was an easy variable to collect data on...* [Reading again from PS4] *Teacher gender, time of day; subject; number of adults in class; class size… You can read the slide…*

38. [Continues.] *Then we go to the second column… Student on-task time…* [Pause (to collect thoughts?).] **T1:** *Neutral teacher-talk time… That is as a percentage… That's basically how much of the observed period a teacher is talking… Overall… Okay? …* [Pause, expecting request for clarification?. Thinking does more need to be said?] **T1:** *So… Number of positive comments about any student's academic performance by a teacher. Number of positive comments about students' social behaviour… That's just student behaviour in the classroom per se. Number of critical comments about academic performance and number of critical comments about behaviour… So… Those are tallies. Just tallies. You can't do those as percentages. Just how many times a teacher has made those types of comments. Not like 'neutral teacher-talk' which is an estimated percentage. None of these can be related to particular students. They are measures of the teachers' verbal behaviour…* **T1:** *The sort of discourse that is being used in the classroom… By the teachers. Just, sort of… A bit of a 'dip-stick' measure… On what's going on… The social-emotional climate for learning that is happening in that classroom at that time…*

39. [Continues.] **R1:** *So… And there is a long venerable research-history to those four tally-categories… Um… Which goes… Particularly, in the midlands interestingly, there was… Um… Erh… There's Merritt and Wheldall… And Houghton. They did a number of pieces of research using those categories… And… And… They came up with a particular set of conclusions from their observations… Which have never really been challenged. Indeed,*

until the publication of my primary study [Apter *et al.*, 2010] *Well… Um…* {**ly** AUT tentative tone: How do I say this tactfully? If I am seen to smear other researchers, this might backfire on current research.} *Well… Their conclusions, which have been published* [replicated] *a number of times. There were just a few surprising details in my study that were different from what they said was the case… And. Um… It's just very interesting… The way academics work.*

40. **A1:** HT **A2: ACT:** [Laughs, knowingly.]

41. **A1:** AUT **R1: ACT:** [Continues.] *Indeed, there was an article today on the radio. About the rigorousness… Or not… Of scientific research… And it is quite interesting… How often… Um… That there is a slippage of convenience… Sometimes. It can be part of the 'cleaning data' procedure. Meaning sometimes… And I am not saying for an instant that previous researchers in this field have done that… That data that is not quite convenient for the hypothesis of the researchers… Is removed…*

42. **A1:** ALL **A2: ACT:** [Chuckle, knowingly.]

43. **A1:** AUT **T1/E1: ACT:** [Continues.] *It is just interesting… That in this* [current] *study…* {**ly** AUT tentative tone.} *I have a rather weird* [anomalous] *result… That doesn't look like the results obtained in any other secondary-school study… Well, in the studies of this particular group of people that keep doing this type of study. The results that I am going to talk about here today… They just don't match up. I'm going to show you…*

44. [Continues.] **T1:** *But it is particular interesting with this current research… I have two datasets. I have the UK sample… And I have a separate dataset that I have got from your school.* [Pause.] *I can make a powerful comparison. It can be sort of used to give us… Added surety… That I am on the right track… In terms of some of the statistical themes… That seem to be emerging. One dataset can be used to test the other. So… Okay…* [Pause.] *Are you all with me so far?*

45. **A1:** ALL **A2: ACT:** *Yes… Mm…* [Etc.]

46. **A1:** AUT **A2: ACT:** [Laughs.] {**Iy** AUT using humour to consolidate enrolment of SLT audience.} *None of you dare say otherwise of course…* [Background laughter.] *So… Right,,,*

47. [Continues.] **E1:** *Let's just look at the first bit of data.* [Clicks PS5.] *In the UK sample… In the one hundred and seven… Er… Lessons. In an average lesson… Say, of about thirty minutes… I found this sort of pattern… Sixteen point eight five positive comments about students' work… Two point nine one comments about students' behaviour… That gives you a ratio of about five to one… Yeh? Do you see how the table works, yeh?*

48. **A1:** ALL **A2: ACT:** *Yes… Mm…* [Etc.]

49. **A1:** AUT **E1: ACT:** [Continues.] *Teachers' negative comments about student' work… About five point four comments in a typical thirty minute lesson…* [Pause.] *About students' behaviour, twelve point three two. That's roughly a one to two ratio.* [Pause.] *Now… That… That five to one ratio... Is high, compared to all the previous research. In previous research, you are talking about three to one and one to three respectively. So this new secondary UK dataset… Interestingly different from Wheldall and Merritt's results… But not at all dissimilar to my previous primary research.*

50. **E1/T2/R1:** [Continues.] *And if you tot up the totals… Positive to negative comments, I get roughly a one to one correspondence with this new UK secondary dataset. So despite exhortations… By psychologists… And other advisers… To teachers… To use three times more praise than criticism in the classroom… Not a lot has changed in thirty years of research…* [Laughs.]

51. **R1:** [Continues.] *Now… Of course… You… Er…* {**Iy** AUT flatters SLT audience that they will have understood the hard data that has been presented.} *As experienced senior teachers… You know this.* [Pause.] *There are no surprises here in anything, I have told you… Right? You are told over and over again to use more praise than criticism,.. Three to one… Right?*

52. **A1:** ALL **A2: ACT:** *Yes… Mm…* [Etc.]

53. **A1:** AUT **R1: ACT:** [Continues.] *But actually, this current research appears to be saying something different that goes against the grain… That is what we are taught… To use more praise… And I am not sure that just because it suggests something different here… That we should throw the baby out with the bathwater… Because actually, there is no huge body of data where teachers… Teachers using very high levels of praise in secondary school… And their effect on on-task behaviour… Are observed… If that makes sense?* [Pause.]

54. **E1/P2:** [Continues.] {**Iy** The next 3 lines counter previous research trends and might appear critical of UK1 school – so AUT is invitational in tone. SLT have hopefully brought along and their support embedded by humour and flattery.} **E1:** *It is important to note here though, that… Primary and secondary schools… Are quite different… Insofar as how students receive praise… For their work… For as you* [especially 'you'] *know, secondary students get embarrassed if you praise them too much… If you are a year nine and you're as 'hard-as-nails'… And you're… You're... You're told that… How wonderful your work is… It can be quite difficult to receive… So… Secondary schools are quite different to primary… Generally, primary kids love to go up in assembly to receive an award from the Head Teacher… A gold star or a certificate or something… So it is different…*

55. **E1:** [Continues.] *So… So as we go on… So here is your school here* [PS6]*… And… And… We've got… You can see straight away… We've got far fewer … Fewer comments… You know… In every box… And let's remember this is year seven through to… Er… Year eleven… It's all the way through… So again… I'm going to try and pull out some of the year… Year by year differences as we go on…*

56. **E1:** [Continues.] *So, overall the verbal behaviour… Let's just… And you can see the ratios and the differences with the national dataset in red… And I really don't know what to make of that… Myself… So here* [PS7] *I've just put in red the UK figure… And the primary figure also… So you see in each case… In each box… The UK figure and the primary figure are roughly similar… Except for… In the bottom right box… Where there is a three point difference… But generally speaking we are in the same ball park… But… Your school… Is quite different. Quite different…*

57. **E1/R1:** [Continues.] *It is almost implausibly low…*

58. **A1:** HT [Murmurs concerned aagreement.] **A2: ACT: {Py** Not sure?} *Mm…*

59. **A1:** AUT [Continues] **E1: ACT: {Iy** This is not good news for UK1 school – lines 59 - 62.} *Compared to the UK sample…* [Pause.] *And you might say that this looks really bad… Teachers in our school don't talk enough and don't make enough comments about either work or behaviour... But, it's more complicated than that… And… And… And… It's those complications that I am interested in and want to ask you about…*

60. **E1:** [Continues] *So here…* [PS8] *Is the neutral teacher-talk… The talk that is neither praising nor criticising… Basically, teaching… Verbal teaching… So in the UK sample we've got fifty two point seven seven… Here, quite a lot less… Thirty one point eight eight… But in the primary sample, we have a lot less again… Twenty eight point three.* [Pause.]

61. **R1:** [Continues.] *One of the things I have been thinking about… Is that… What happens in the classic primary lesson… Is that you get the exposition at the start of the lesson… The*

teacher talks for quite a considerable time… Then, 'Get on with your work' and then usually there's… Round tables, group work. A bit of buzz and noise. In maths… Whatever subject… And you get on with your work and the teacher circulates…

62. **R1:** [Continues.] *So there is a bit of a teacher-chat at the start and then the teacher circulates making odd comments and helping individuals… Now I think that secondary is a bit different to that? … The teacher is involved in more whole class talking throughout the lesson… That's my impression… I don't know… I want you to explain that… I don't know…* [Pause. Nobody speaks.] So… *We'll come back to some of this… If you want…*

63. **E1/P2:** [Continues.] {**Iy** But the good news is… Or is it good news?} **E1/P1:** *So… So… So this is lovely…* [PS9] *Your kids are no less on-task than the UK sample… More or less… It's an insignificant difference… So they're working… Your kids are working as hard as the UK sample… Which is… Again… If… If… The work rate is related to teachers working hard verbally… Which was my previous hypothesis with the primary research… Then that shouldn't be the case…*

64. **A1:** ALL [Murmur agreement.] **A2: ACT:** *Mm… Yes…*

65. **A1:** AUT [Continues] **A2: ACT:** *But it is… You kids seem to be working as hard…*

66. **A1:** HT **A2: ACT:** [Agreeing.] {**Py>Iy** Tentatively picking up on the bad news. Implicitly: but what does it mean?} *But there is less teacher involvement…*

67. **A1:** AUT [Continues] **E1/P1: ACT: E1:** *But there seems to be less teacher involvement…* {**Py>Iy** Tentatively picking up on HT's response to bad news. Implicitly: but what does it mean?} *I don't know… You tell me… This is it… This is my headache. So… So… Just comparing the... The… The… On-task means… Here…* [PS10] *Now this is interesting in your school… I just picked out year seven… So on-task mean throughout the school, seventy-eight point oh nine… But year seven has a significantly lower on-task mean… Significantly lower at sixty two point eight nine… And yet the primary key stage one two mean is eighty five… So it would appear… Your kids come in to the school, very undisciplined and not wanting to get on-task…* {**Iy** This <u>is</u> good news.} *And then they improve massively over that year…*

68. **P1:** [Continues.] *That's what appears to happen… You do something very well in year seven… That compels your year sevens to learn to get on task… 'Taming them', as it were…* [Pause.]

69. **P1:** [Continues.] *And there could be an issue with some of your feeder primary schools maybe? … I don't know… I don't know your feeder primaries well-enough really… But…*

70. **A1:** AH1 **E1: ACT:** {**Py** AH1 apparently buying-in to good news.} *So they are improving through years eight and nine then?*

71. **A1:** AUT **E1: ACT:** *I'll show you that…*

72. **A1:** AH1 **R2: ACT:** [Continues.] *I'm wondering… How quickly… You use the word 'taming'…* {**Py>Iy** AH1 damns AUT with feint praise.} *Which I find amusing… I know what you mean though… Did you find that… I'm interested in how quickly it happens…* [Pause.]

73. **A1:** AUT **E1: ACT: {Py** Sensing mounting attack from AH1} *Yeh, I'll show you that…* [PS11] *Can you see that… Yes, you can see that… Here is the UK sample… And here is your school… The top line is the on-task behaviour… And its not a very smooth curve… But you wouldn't expect it with this sort of statistics… This sort of graph… But you have a trend in the UK sample… For the on-task rate to go down year by year… But your trend,,, In year seven… Is that steep rise in year seven… But a general trend upwards… Upwards towards the upper years…*

74. **E1:** [Continues.] *You can see that teacher input diminishes in both samples… The UK sample and your school…*

75. **A1:** HT **E1: ACT:** *Yes… You can see that is very similar… What is… What is the green one, Brian* [AUT] *?*

76. **A1:** AUT **A2: ACT:** *The green line is neutral teacher-talk… Verbal teaching.*

77. **A1:** HT **A2: ACT: {Py>Iy** Sensing mounting attack from AH1 and forming alliance with AUT.} *That's very interesting…*

78. **A1:** AUT **A2: ACT:** *Again the curve is not smooth… It's a little bit bumpy… But the trend seems consistent.* [Pause.]

79. **A1:** AH2 **T1/R2: ACT:** {**Iy** AH2 (unexpectedly AH2) uses an extended series of questions to wrest authority from AUT and attack the research.} *Like… You are talking about a UK sample… And you are talking about lesson observations here… Predominantly… What were the lesson observations here? Was it the whole range? Was it top middle and bottom ability? What was the starting point with the UK sample? Where did the students start from? What were the abilities of the students? How many lessons in a good school? What was the profile of the teaching like… In the UK sample? Were they NQTs or were they more experienced? Etcetera… Surely those variables… I know that you are looking at… You are looking at… These predictions that you are making… There are so many variables… Even down to the fact that where the students were on the curriculum… What was being expected of them at that time… You know… Did they have a test coming up? Is that why the teacher talked more? Is that why the teacher talked less? How much does it take into consideration those variables? Because those could be some of the reasons behind some of the data that you've got.*

80. **A1:** AUT **T1/R2: ACT:** {**Py** Partial agreement, AUT gaining time to consider what was the subtext. It felt like an attack.} *I didn't control for the 'ability' variable, because ability grouping was not something that I was interested in… Overall, I have two large datasets… The variables that I am not interested in should largely average out in each dataset as long as the samples are properly representative of the variables that I am interested in… Particularly, on-task behaviour and… And… Teachers' verbal behaviour. I have factored in most of the other variables that you listed… I think…*

81. **A1:** AH2 **R2: ACT: {Py>Iy}** *What about times of the day?*

82. **A1:** AUT **T1/E1: ACT:** {**Py** A firm volleyed return; AUT becomes clear, and re-establishes authority.}*Yes… I have two variables that factor in time of the day… The time… And also whether the lesson before or after lunch… I am interested in those variables… Interestingly, but in one way disappointingly… I didn't find anything particularly interesting with either dataset in respect of the time of the day… Which surprised me… With the primary research, I expected to but didn't… And the same with the current research.*

83. **T1/R1:** [Continues.] *Interestingly, there is an aspect of this time of day thing that is difficult to control for… Some school now have a really early lunch time. And a much longer afternoon…*

84. **A1:** ALL **A2: ACT:** [Murmurs of agreed understanding.] *Mm… Mm…*

85. **A1:** AUT **R1/R2: ACT:** *What effect this aspect of 'time of day' might have on on-task time, I shall have to cover in the discussion section of my thesis… But overall… Whilst you and I might have predicted otherwise, there is no statistically significant association between time of day and on-task behaviour… The more interesting variable effects are coming up shortly…*

86. **A1:** AUT **E1: ACT:** [Pause.] *Okay…* [PS12] {**Iy** AUT resumes standpoint in discourse and position in meeting.} *With on-task behaviour… The strongest correlation… And I have*

checked this with a range of statistical techniques… The strongest correlation… And the strongest predictive or causal association that I found in both datasets… Is between the critical comments used by teachers about student social behaviour and the 'on-task' behaviour variable… Praise doesn't come into it... It is the teachers who jump early and hard on anti-social behaviour who achieve the highest rates of on-task behaviour in their classes…

87. **E1:** [Continues. Re-emphasising.] *So that is significantly different in secondary schools in comparison to primary schools… In secondary schools… Where a teacher is on it… Where they are picking up on it… And commenting… And not letting it go… You are likely to get more on-task behaviour.*

88. **E1:** [Continues.] *And that is correlationally true… It is also causally true, using an analysis of regression statistic. How ever I did it… Assuming parametric or non-parametric data distribution… This was still the dominant relationship.* [Using different language to re-emphasise the point.] *A teacher in the secondary context relentlessly picking up on bad behaviour will get more on-task behaviour…*

89. **R1/P2:** [Continues.] {**Iy** AUT able to be magnanimously deprecating about psychologists and the psychological standpoint – a paradoxically status-enhancing device.}*That is contrary to most of what psychologists have been advising teachers… That praising kids is more effective than criticising them… And those psychologists include the researchers that I previously quoted today…* [Pause.] And yet… *And yet this is a common-sense view and… Er… A tabloid-shared view… A 'strict' teacher… Is a good teacher?*

90. **A1:** ALL **A2: ACT:** [Murmurs of agreement.] *Mm… Mm…*

91. **A1:** AUT **E1/P2: ACT:** {**Iy** AUT develops standpoint, lines 91 and 92, in discourse and further enhances position in meeting.} *In terms of associations between variables that are correlations… But not causal… Year group is correlationally associated with on-task behaviour… Weakly with the UK sample… But more significantly with the dataset that is derived from your school. Meaning, the higher you go up in the school, the more on-task behaviour they are… So… Smaller groups… More specialised teaching… It appears that they are going to be much more on-task…*

92. **E1/P2:** [Continues.] *And this one* [PS12; bullet point 3] *I hoped to be able to show, I guess… Correlationally… Both the UK sample and the one from your school show… The more years of teaching a teacher has under their belt… The more experienced they are… The more on-task their students are… More so than for NQTs..* [Newly qualified teachers] *It would seem…* [Pause. Waiting for a reaction.] *I am looking at you* [AH2] *now…*

93. **A1:** AH2 **E1: ACT:** {**Py** Sounding conciliatory.} *No, I am just, like, correlating what you are saying with some of the results and outcomes that we get then…*

94. **A1:** HT **A2: ACT:** *Yeh, yeh. Yes…* [Knows what AH2 is going to say and agrees with AH2's query.]

95. **A1:** AH2 **E1/P1:** [Continuing.] *…So if our students are compliant… Even moreso than the UK sample and standards…* {**Py** Valuing of AUT's message – and now attempting a synthesis between his own position and AUT's standpoint.} *We should be getting better results than what we are, and it throws up the question of… What our school's adviser Mick Waters has been banging on about all the time… 'Stretch and challenge'…* [Meaning: AH2 believes they not stretching and challenging students enough.]

96. **A1:** HT **A2:** [Interjecting.] *Yeh…*

97. **A1:** AH2 **P2:** [Continuing.] *And are we stretching them enough…*

98. **A1:** AUT **E1: ACT:** [Pause.] *Well… Okay…* {**Py>Iy** Agreeing to synthesis between AH2's position and his own standpoint. Resuming authoritative reporting of the results.}

There are some caveats to a lot of this… And I shall show you what they are… [PS13] *So there are two missing correlations I expected to find, and which I haven't… There is an insignificant correlation between how much a teacher talks and student time on-task. So, in the primary sample I have a positive correlation. Here, in the UK sample and in. In. In. Your school… I don't find… So that doesn't seem to be significantly correlated… There is a very slight effect, but it's not reportable. Um…*

99. **E1:** [Continues.] *And also… Academic praise… There's a very slight… There's a weak correlation… Between praise and on-task time… In your school… But I can't report it confidently. In the UK sample, there's no correlation at all.*

100. **P2/E2:** [Continues.] *But I have to factor in the factor also… That overall… I am now expecting secondary phase kids… To be on-task at a lower level than primary school kids…* {**Iy** AUT is surprised that this does not have a significant effect on SLT. AUT attempts to emphasise the point.} *Overall… Year by year…*

101. **E1:** [Continues. PS14] *Now because of your concerns…*{**Iy** New subject. AUT develops the case for relevance to the UK1 SLT in respect of authoritatively drilling down in to the maths and English statistics.} *In your school… About maths, and your celebration of English… And how things were working… At that time…* [Immediately, post-Ofsted inspection in 2014, when maths had been cited as a particular area of concern.] *I took a particular interest in those two subjects… So students were significantly more on-task…* {**Iy** AUT emphasises his informed familiarity with UK1 school's circumstances.}*So this is all about your school now… In English… Throughout… Throughout… In all years. Eighty three point three seven… Which is almost up to the primary on-task figure… And… Than in maths… Which was seventy one point one eight… Remember that is two thousand and fourteen and I know that there have been lots of changes since then…* [Major reorganisation of the maths department and a number of staff 'let-go'.]

102. **E1:** [Continues. Reading from PS14.] *So English lessons were characterised by higher levels of positive comments about students' work and higher levels of criticism of behaviour…*

103. **E1:** [Continues. Reading from PS14.] *Maths lessons were characterised by similar levels of positive comments about students' work but comparatively higher levels of criticism of behaviour… So basically… Maths teachers were working harder… Criticising bad behaviour… Which follows, if maths lessons aren't being successful…*

104. **A1:** AH2 **A2: ACT:** {**Py** AH2 is compliant.} *It makes sense…*

[From this point onwards in the meeting the dialogue between AUT and ALL became much more excited and faster in pace, with more rapid conversational turn-taking. {**Py** There was also an evident reduction in expressed scepticism: from the sceptical position represented by AH2, line 79 (above).}]

105. **A1:** AUT **E1: ACT:** [PS15. Pause.] *So this is year by year…* {**Iy** Invitational.} *But I've got this strange dip in year eight… In English…* [Pause.] {**Iy** Invitational but showing collaborative leadership style.} *Anybody? Any ideas?* [Pause.]

106. **A1:** CT1 **E1/P2: ACT: {Py>Iy** CT1 responds to invitation by sharing it with SLT group.} *That's the current year ten, isn't it… Which…* [Unison with HT.] *Makes a lot of sense…*

107. **A1:** HT **E1/P2: ACT:** *Yes…* [{**Py** Highest status SLT member (HT) joins with lowest status (CT1) and model working together.} Unison with CT1.] *Makes a lot of sense…* [HT and CT1 in strong agreement?]

108. **A1:** ALL **E1/P2: ACT:** [{**Py** Everybody joins in – SLT group compliance with HT.} Loud laughter. Joint recognition of a problematic rogue year group?]

109. **A1:** HT **P1: ACT:** {**Iy** HT develops subsidiary hypothesis.} *And also it would have been interesting to see who was teaching them* [English] *when they were in year eight… To see if that is the reason…*

110. **A1:** AUT **R2: ACT:** *Yes…* {**Py** Does not want collude by identifying an apparently 'failing' teacher.} *Obviously that would be up to you to find out rather than me* [because of confidentiality / anonymity agreement with individual teachers.] …

111. **A1:** HT **A2: ACT:** *Yes, of course… But this is really interesting… Can I ask something?*

112. **A1:** AUT **E1: ACT:** [Reason for AUT ignoring question from HT: apparent follow-up on 'failing' teacher.] **{Py** AUT and SLT beginning to join together in enthusiasm.} *But the maths… Maths…* [Over excited hubbub in group.] *The maths curve… You can see it makes quite a lot of sense…* [Noting the slump in on-task behaviour in year 10 / 11: GCSE exam year.]

113. **A1:** HT **A2: ACT:** [Slow thoughtful agreement.] *Yeh… Yeh… Yeh…*

114. **A1:** AH1 **R1: ACT:** [Pause.] {**Iy** AH1 tentative invitational of alternative hypothesis.} *But we used to say about English and progress though… You sort of start… With a gentle progression… Then hit a rapid progression at the end of key stage three… And you carry it through…*

115. **A1:** AUT **P2: ACT:** [Pause.] *So… Okay…* [PS16] {**Iy** AUT ignores AH1 and endeavours to drive the discussion forward to primary hypothesis.} *So what's the hypothesis… The hypothesis that this is working on, is that: learning in the academic classroom is…* [Pause.] *In-part… A function of the social-emotional climate for learning.* [Reads from PS16.] *That climate can be characterised, measured and quantified by examining the way that teachers talk and feedback in the classroom and how this effects the time that students stay on-task…E. G.… A good case study… Your English… Your English… And talking to Anne* [HT]*… About English and how it's been working… Um… Has sort of… Sort of… Reinforced that view… Your English teachers create a Social-Emotional Climate in… In… Their classrooms… That is particularly conducive for learning…*

116. **P2:** [Continues.] *Now…* {**Iy** AUT continuing to drive the discussion forward using other research to support hypothesis.} *Do you know John Hattie's research* [Hattie, 2008]*… About effects sizes… Good… Because I've bought an overview here to leave with you… You might wish to copy this and distribute it to teachers… Whoever you want to… What Hattie said… He measured the effect sizes of different interventions in the classroom… He used a meta-analyses of a large number of studies… He said that the effect size of… Teacher feedback… Was the largest effect size of all at one point one three… What does that mean? One point zero is equivalent to a two grad leap at GCSE, he explains…*

117.**R2:** [Continues. Pause.] {**Py** AUT reintroduces contradictory view in lines 117 to 119, as per AH2, line 79.} *However, there are other factors that my study has not examined at all… Some of the variables that you* [to AH2] *were asking about… The study does not address the curriculum; what is actually taught; the materials a teacher uses to teach; or their teaching techniques; the techniques they use to explain… Or embed… New knowledge….*

118. **R2:** [Continues.] *And the cultural baseline… And the cultural baseline… Hattie recognises that as a significant variable… Two grades again… So year sevens are coming in… Generally speaking… Not wanting to… In any way… Besmirch this wonderful city… But we aren't blessed with a majority of encultured and wonderfully affluent catchments… I think you know what I am saying…*

119. **R2:** [Continues.] *And my own understanding of that… Is that this school has a large number of C. D. P. D. students… Culturally deprived and parentally disadvantaged… We do seem to have here, a large number of student who might be described in this way… And, and… Hattie is saying… That there must be a huge effect from that in your school… To start with… You know… So whilst we can make progress, it will always be an upward struggle from such a baseline.*

120. **P1:** [Continues.] *And these are some of the questions* [PS17] *that I think we might address now… Rather than read out these... Can I ask you to read them for yourselves? …*

[Long pause – 2 minutes.]

121. **P1:** [Continues.] {**Iy** Offers limited democratic choice to SLT.} *I'm not sure what you want to do with these* [questions]*… Shall we take them one at a time?*

122. **A1:** HT **A2: ACT:** *Yes. I think so…*

123. **A1:** AUT **P2: ACT: {Iy** AUT invitational but maintaining control.} *Shall we go with one…* [Pause.] *What do we think? So… If this… So… If we are saying that your teachers are not being… Well there are two ways of looking at this… If teachers are not being as verbally… Er. As verbally vigorous… As they could be… Either… Well, either there is an observational effect… And that is not reflected in the actual level of verbal vigour or it's not… But either way… Is there a way of investing people with more enthusiasm… For more talking in the classroom… Active teaching…*

124. **A1:** HT **R2: ACT: {Py>Iy** HT accepts invitation and takes control.} *I think there is a shift, Brian* [AUT]*, over the last couple of years… It used to be… It was all about… The performance of the lesson… And it was all about the teacher… Being all singing, all dancing… And it was all about activity and it was about students up and about doing things… In twenty fourteen, the shift started away from that to teachers just being facilitators, far more engaging the children, getting the children to take charge… So some staff have been taking, sort of, a more deliberate back seat… A planned approach to their teaching and learning to let children make their own way… So whether there is any link between the styles of how people deliver…*

125. **A1:** AUT **P2: ACT:** {**Py** AUT knocked back.} *Well sort of… Roll that in to question four…. Because… Maybe there is a difference that can be made… One of the things that I have been thinking about…* {**Iy** AUT responds by reintroduces issue of Year 7 cohort's under-developed baseline and the solution another school has attempted.} *Is that if your year seven on-task baseline is low… As the students arrive in your school … Maybe that's…* {**Py** Responds to own question. Offers apparently new solution which is outside remit of current research.} *We've talked about another secondary school a couple of miles away, where they operate a primary model… In the bottom part of the school… Where kids have the same teacher for most subjects and don't move around much… And why they do that… They… They… Their view is that… That they… These kids need more nurturance… Than a more usual secondary model would afford them… Because the kids need to reach some sort of emotional maturity… First, before they can… I don't know…*

126. **A1:** AH3 **R2: ACT:** {**Py** Partial but modified agreement.} *I don't know… I think the more that we get a handle on our baseline… And what our feeder primaries are doing… The more we will be clear on what skills our students are coming through with… I think you will see a difference in that… I think that generally in the past, we have taken a step back and re-taught stuff that had already been done in primary… And therefore it takes a dip before it comes out of that trough and starts building again…*

127. **A1:** AUT **A2: ACT:** *Yes…*

128. **A1:** AH3 **P1: ACT:** *And it'll be interesting to see… The recent work we have done with baselines and feeder primaries…* {**Iy** Continues to develop solutions outside of scope

of research.} *It's getting departments to realise what their curriculums involve and what skills students come with… Therefore we should hit that ground more like running… And we might then see… It might not be steep… But more of a gradual incline… And to carry it through to years eight and nine…*

129. **R1/P2:** [Continues.] *Our students come across as very receptive… Even… Perhaps, the teaching is not good in some areas… Ours students are still receptive and patient…* {**Py** Partial but modified agreement.} *It's interesting when we talk about… Er… Verbal feedback… And generally the feedback as a whole across the school… Is that whether we are challenging enough or supportive and pushing forward enough… And higher attainers… Just seeing students are on-task or not and whether we are happy* [complacent] *with that or not…*

130. [Continues.] *I think praise is a very complex thing, isn't it…*

131. **A1:** AUT **A2: ACT:** *Yes…*

132. **A1:** AH3 **R2: ACT:** [Continues.] *And I think it is very easy to put all your eggs in to one or two baskets… And to presume eight, nine hundred people will respond to two strategic positive approaches to praise… Because I think it is a lot more complex than that…*

133. **A1:** AUT **A2: ACT:** *I agree…*

134. **A1:** AH3 **R2: ACT:** [Continues.] *And I think you may see a difference if you really thought through the way in which you praise… The way in which you rewarded students… In a much more complex way… If you knew much more about them… And the way in which you could make…* {**Iy** Continues to develop solutions outside of scope of research.} *To me… What stands out… Is that there is much more work to be done around… Like… Fostering this love of learning… And the more of that you can do… The more inquisitive the students are… The on-task behaviour will change… Because the intrigue and the interest is there… I think they are pretty complex things, you know… To be able to…*

135. **A1:** AUT **A2: ACT: {Py** Apparent compliance.} *I agree… I agree…*

136. **A1:** AH3 **P2: ACT:** [Continues.] … *Crack it… And I think it would take intelligent people and some real… Like…* [**Iy** Implied criticism of HT?} *Clear leadership… On how you are going to do that…*

137. **A1:** AUT **A2: ACT:** *Yeh…*

138. **A1:** AH3 **P2: ACT:** [Continues.] … {**Iy** Continues to develop solutions outside of scope of research.} *And I know that there are models out there… The Opening Minds Curriculum… That type of stuff… But how brave are you and bold… Especially when you are in a school, that's been judged inadequate previously… And you are taking big risks with… You know… Because you've got to explicitly show progress… And that's where I think some of the teacher behaviour comes* [**A1:** HT **A2: ACT:** *Yeh, yeh…*] *from… In my opinion… Is the paranoia of… It's all about progress, progress, progress* [**A1:** AUT **A2: ACT:** *Yeh, yeh, yeh…*] . *And sod the rest… Because if I can show the progress…* [**A1:** HT+AUT **A2: ACT:** *Yeh, yeh…*] *No one can argue with me… And I don't think… Going back, to what you said at the start…* [**A1:** AUT **A2: ACT:** *Yeh.*] *That things like Ofsted help… Do they?* [**A1:** AUT **A2: ACT:** *No. no, no.*]

139. **A1:** HT **A2: ACT: {Py** Distancing and defending from implied criticism.} *People like that…*

140. **A1:** AH3 **A2: ACT:** [Continues.] **{Py** Admission.} *That's where… That's where the crux of some of the problem is… And I'm guilty of that… I sent out a massive email today… About Ofsted… These are the things that you do* [adding to colleagues stress]…

141. **A1:** AUT **T2: ACT:** *I know…* **{Iy** The reality: AUT demonstrates empathy.} *I hear from Anne that you've probably got your first Ofsted as an academy… Probably Easter…*

142. **A1:** HT **T2: ACT: {Py** Accepts the conch of leadership back from AUT for final part of meeting.} *Easter is earliest that they can inspect us as a new school… Chances are, and we'd hope… That it's going to be a bit longer… Because we'd hope to embed our changes a little bit more…*[**A1:** AUT **A2: ACT:** *Yeh, yeh.*] *For instance, the whole approach to key stage three… From a leadership perspective, from our knowledge of what is going on… It's a lot tighter… The data and the understanding of how children are progressing… Is much tighter… Since John and Chris* [first names AH1 and AH2] *have been looking at that… And also the fact that there are now no levels in the national curriculum… It means that if you came into key stage three now you'd see different approaches because it is very different to what it was…*

143. **A1:** AUT **A2: ACT:** *Yeh, we discussed that, I recall…*

144. **A1:** HT **R1: ACT:** [Continues.] *But it's useful though because our English data on its own entirely confirms your data… Because last year… Those children… Who would have been part of that… Achieved the best set of results we've had… And maths have improved but… It wasn't yet…* {**Iy** Continues to develop solutions outside of scope of research.} *I'd be really interested to see if the maths department now, and a different staff and a different belief that they can do as well as English eventually… What the change would be… Because children do...* [Pause.] Now*... Enjoy going to maths… Which they didn't… They didn't a couple of years ago… There's that sort of link…*

145. **R1:** [Continues.] {**Iy** Continues to develop solutions outside of scope of research.}*And I also think that there's further study needed, looking at the style of English teachers who… Um… Communicate the love of their subject… And are good at expressing positivity… Perhaps a little bit more easily than some traditional maths teachers… Who are very narrow in their social and emotional intelligence… About people…* [**A1:** AUT **A2: ACT: {Py** Partial agreement.} **[**Chuckling. Appreciative.]] *But who are better in their… But you know what I mean by that…*

146. **A1:** AH2 **A2: ACT:** *Yeh, exactly…*

147. **A1:** HT **R1: ACT:** [Continues.] {**Py** Responds to self-provocation by developing argument.} *But they are… Some of them who are good mathematicians don't know how to communicate with adults… Never mind children... They can impart knowledge, and if you've got a love for maths… And you're good at, maths... Then that… If you broke down maths and you looked at, say, the year twelve and the year thirteen…* [**A1:** AUT **A2: ACT:** *Mm…*] *Mathematicians* [students] *and what their responses are… It would be a very interesting… Different picture to say bottom set year nine… With a certain teacher…*

148. **R1:** [Continues.] {**Py** Returns to confirmatory benefit of current research. Magnanimous leadership equals authoritative standpoint.}*I think that there are some nuances to come out within the different subjects... But absolutely, that* [UK1 dataset and AUT's analysis] *confirms absolutely what we believe…* [**A1:** AUT **A2: ACT:** *Yeh.*] *It's better than we would have thought… Based on what was written about us two years ago… But absolutely we need to move and look at what is different now… Based on… All the things that have changed since you were here last…* [**A1:** AUT **A2: ACT:** *Yeh.*]

149. **A1:** AH3 **A2: ACT:** {**Iy** Moving back from current research.} *You're delving in to the realms of… Like… What's that guy's name? Bob Hymer?*

150. **A1:** HT **A2: ACT:** *Barry Hymer* [Hymer,Whitehead and Huxtable, 2008]…

151. **A1:** AH3 **R1: ACT:** *And he touched on this, didn't he?* [**A1:** HT **A2: ACT:** *Yeh…*] *He said exactly what you're saying… You know, this drive for the recruitment of teachers… Is based around academic qualifications… When really, the prerequisites are to be a good*

teacher… And the level you've got to teach at… How much is it based upon… Academic… Like… Intellect… Or how much is it based upon the natural attributes you've got... And I think there is some work around that... And I think that's why this school is partly successful… [**A1:** HT **A2: ACT:** *Yeh, yeh…*] *Is the things you've done and the leadership team before me… Have brought in people… Yes, they're not qualified teachers… But… You know… Oh yeh…* [**A1:** ALL **A2: ACT:** [Shared laughter – in recognition that CT1 was taken on unqualified and was now a successful member of the SLT.]

152. **A1:** HT **P2: ACT:** {**Iy** Partial acknowledgment of UT's research. But HT leadership will involve hiring and firing.} *Yeh, but it's about… It's the right people… People who are going to engage the young people that we've got… And it's about being sure…* [**A1:** AUT **A2: ACT:** *Yeh.*] *That we've got… Across the whole profession for teachers… Whether they are qualified teachers or instructors… Or unqualified teachers… That if somebody's got a love of a subject and can impart knowledge to a child… We can do the rest as a leadership team… We can up-skill them… Give them opportunities to do research… And we can get them in… Put them with children that they'd probably never want to work with again* [rye tone]*… Get them into the teaching and get them qualified and move them on in their career…*

153. **R2:** [Continues.] {**Iy** Moves back to acknowledging the role of feedback and the UK1 dataset.} *But the bit about… Just as human beings… The link between the number of positive statements in a meeting or a lesson, against the criticisms… We haven't got right… We haven't …*

154. **A1:** AUT **A2: ACT:** {**Py** Checking in. Are HT and AUT on the same side.} *You don't think that you've got that right…*

155. **A1:** HT **T1: ACT:** {**Iy** We acknowledge the research but we are heading in a different direction.} *No… No… Our shift at the moment… I got the staff to do their own… I called it 'three to one'… When I came back* [after a period of illness] *I wanted to meet with every member of staff… I'd been off for a few weeks… No way could I get to one hundred and twenty staff… So I wanted to know how they were feeling… So I got them to jot down three things that were doing really well… Or they're proud about… And one thing… They could… If they could change… Or I could change for them that would make things better… And I had pretty much everybody sending something back… Non-teaching staff… Social staff… Everybody… And it's really interesting about saying… Look… People find it really hard to put three positives down about themselves… They could easily make a list of negatives…*

156. **A2:** [Continues.] *Like… If you approach your teaching like that with kids… And we all know if you taught in other challenging schools… It doesn't matter what's happened the day before… You've got to go in with your smile…* [**A1:** AUT **A2: ACT:** *Yeh.*] *… And when they say things that are completely outrageous and wrong and misplaced… It's how you do the… Oooh… Well done… Not quite right but have another go… But also… If you read the latest stuff from Carol Dweck, actually… They don't want you to do that… They want you to say… No, that's wrong… Kids will accept that from you and kids will not expect you to flannel it…* [**A1:** AUT **A2: ACT:** *Yeh.*] *There's a lot of work to be done with just the simple… Adults… And how they converse with children… In school…* [**A1:** AUT **A2: ACT:** *Yeh.*]

157. **A1:** AH1 **R2: ACT:** *So it's interesting as well… Is what we deem is good progress… And how we quantify progress… Because a lot… Especially with our school… There were elements of that… For instance, the amount of written work… They were looking at that for evidence that that is progress…* {**Iy** We acknowledge the research.}*Rather than when we are talking verbally… The teachers are talking verbally… How is it… How is it… Er… Measured as the student verbal feedback?* [**A1:** AUT **A2: ACT:** *Er… Yeh.*] *In pushing their progress forward… In the verbal context…*

158. **A1:** AUT **A2: ACT: {Py** Appealing for further agreement with central hypothesis of research.} *I was… I was intrigued… And I sort of… The social emotional climate for learning* **[**Anderson and Brewer (1945; 1946a; 1946b)] *… Is a term which I have coined… It's not SEAL…* [Social Emotional Aspects of Learning (Humphrey, *et al.*, 2010)] *Though parts of SEAL were good… It was from the nineteen forties… It was a study in nineteen forty five… And what intrigued me was the move away from viewing that… The climate of the classroom as conducive to learning… To…*

159. **R1/R2:** [Continues.] *You get to the eighties… Ted Wragg's books… And he pulls out the idea… No longer are we worried about this constructed universe, the classroom… The social-emotional climate for learning… We are now worried about teachers' performance… That to me is an interesting shift… In the whole way we look at what goes on in the classroom… Teachers' performance… Teachers' performance… Teachers' performance…* [Repeats x3]

160. **R1:** [Continues.] {**Iy** Recontextualising discussion in to the current research.} *But there's something much more complex… Much more complex things going on… Praise… You know… Anne said that we haven't always got this right… And you* [to AH3] *said praise is much more complex… Praise… You know… You* [to HT] *gave a lovely example… You said, you haven't got it quite right but well done… That's maybe talking to a year nine… Maybe… But a year seven… Who's come in with loads of failure… You might just say, well done! End of. Do you know what I mean?* [**A1:** HT **A2: ACT:** *Yeh.*] *Then you might make a mental note and say the next day you are going to meet up with that year seven and say, next time try this… And you leave it a day, because that year seven can't assimilate the two parts, the praise and the 'but'… It is at least that complex…*

161. **A1:** AH2 **R2: ACT: {Py** Partial agreement.} *We do that with teachers…* [**A1:** AUT **A2: ACT:** *Yeh, yeh!*] *We do that with the people we are coaching and mentoring… It's really the stage they are at in their career…* [**A1:** HT **A2: ACT:** *Yeh.* [Laughs, appreciatively.]] *I did that yesterday…*

162. **A1:** AUT **A2: ACT:** [Lengthy metaphor about the difficulty of getting a person to change their behaviour.] *The only other… The slight… Metaphorically… Who knows where their car keys are in the morning? When you are going out to work… Who puts their car keys in the same place?I was talking to a psychiatrist and he was talking about his wife… Who is also a psychiatrist… And he said he'd tried to get her… Many times… To get her to put her car keys in one place… And he said that every morning they'd have a drama… She'd say… Where's my car keys? Where's my car keys? Where's my car keys?* [**A1:** ALL **A2: ACT:** [Laughter of recognition.]] *And he said…In the end, I got fed up with this… So I now say… Your car keys are your business… And still every morning she says… Where's my car keys? I've put hooks by the back door… He said… For her to put the car keys on… I have put a bowl for her car keys by the front door… I have tried everything… But still she won't learn.*

163. **A2:** [Continues.] {**Iy** Fighting back. Bringing the meeting-focus back to focus of current research.} *When you spoke about a maths teacher…I thought of the psychiatrist's wife who found it difficult to learn… It is very hard… When you know… As a teacher… You know… Why I have given you the Hattie document… Is because… Maybe it is something that you give to every blooming teacher in your school…* [**A1:** HT **A2: ACT:** *Yeh.*] *Because… Feedback is the key… To optimisng the social emotional climate of the classroom for learning… Feedback… And yet… Every teacher knows that… But many teachers don't do it… And yet… This learning would make their lives in the classroom… So much easier...* [**A1:** HT **A2: ACT:** *Mm.*]

164. **R1:** [Continues.] *And you think… When you mentor and coach teachers… Why don't they do it differently? It would be so much easier for you… If you did this, that or the other… And like the psychiatrist you think I am getting nowhere with this…* [**A1:** ALL **A2:**

ACT: [Laughter of recognition.]] *This is not changing… And this is a big task in your school… To get people to… Because most teachers would say… That they know all of this… So why don't they do it?* [**A1:** HT **A2: ACT:** *They forget...*] *Here with this research… In both samples… You've got one to one... Positive to negative.*

165. **A1:** CT **E1/R1: ACT:** {**Iy** CT supports research by expanding on indirect feedback technique.} *It is interesting… Today… We've been doing about obedience in psychology* [as part of CT's teacher-training course]*… As you know… This idea of the social-behavioural negatives… I used some of the AP:BiS stuff…* [Applied Psychology: Behaviour in Schools in-service course (Apter, 2004).] *… Proximal praise… I used the example of… If somebody is not sitting up correctly… I used as an example in a lesson… And one of the lads said… Sir, you know I almost felt like sitting up then… I never said anything negative… I'd just go in the room and say… Well done for sitting up straight… It was bizarrely out of context… But you don't necessarily have to go on at a student… Or tell them off… To get the appropriate result out of them… So that's is a powerful way to shift that ratio I think… Putting it simply.*

166. **A1:** AUT **P2: ACT:** *Yeh… The old way of doing things is to say to a school… Here's this course… Here's this training… Assertive discipline… Or whatever it is… This is the way to do it… But all I am suggesting to you, the senior leadership team of this school… I am saying this is what it looks like… Simply… An interesting picture of your school… Do something about it… Design something yourselves that precisely addresses this situation… I've given you the Hattie document as a catalyst…*

167. **R2:** [Continues.] {**Py>Iy** Responding to / checking out perceived scepticism.} *Look… We're at five now… I am sure that you are ready to go home and have some tea… But I hope that this has been useful…* [**A1:** ALL **A2: ACT:** [Murmurs and yehs of polite agreement.]] *Well… To be honest… You are not going to say here and now that it was rubbish….* [**A1:** ALL **A2: ACT:** [Laughter of recognition.]]

168. **A1:** HT **T2/R2: ACT:** {**Py** Reassurance.} *No, not at all… We had our professional conference last week… Every year we go… The whole leadership team… It starts on a Friday and we have an INSET* [In-service training] *day… So staff do their moderation… And we go off and get inspired about why we do what we do… Because you can get very bogged down obviously with the negatives… Which are often three to one the other way…* [**A1:** ALL **A2: ACT:** [Laughter of recognition.]] *And so we listen to all sorts of people… Some who are brilliant… Some who are not so good... And then you come back and you feel energised about what it is you do… And why you do it… And you get to think differently from the mad treadmill that sometimes you can't get off… This… This is totally appropriate on the backs of some of the conversations we've had…*

169. **R1:** [Continues.] *And I have been thinking about how we shift… Not just because of Ofsted coming in… But that is driving what we're doing… It's bonkers of course… But it has to… If it didn't… It would be a brave or foolish Headteacher who didn't do what Ofsted said regardless… That's built in… But we are doing it because it is good. We are shifting significantly… The approaches… The leadership team… The teaching… We are changing everything about the culture of this school… It is already very different from what it was…*

170. **R2:** [Continues.] {**Py** Reassurance.} *Your research confirms some of the stuff that we know… And also makes us think… It would be interesting to see your study done again with our key stage three now…*

171. **P2: A1:** AUT **A2: ACT:** *Well, yes… You could organise that… You've got the tools to do that yourselves… Using teacher-peers in a non-threatening way… I have always treasured the hope that a school would adopt a peer-teaching monitoring and measuring model… One teacher would say to another… Come in to my lesson and… You've got a bit of time… I'll do the same for you next week… Come into my lesson with one of them*

MICRO measuring things… And tell me how I am doing… I'll be as normal and as natural as I can be whilst you are watching me… Tell me what I am doing verbally… What the discourse is like in this lesson with this particular group… I'd love teachers to do that… But clearly they are not going to do that if they feel stressed and defensive… With Ofsted coming…

172. **A1:** HT **R2: ACT:** *Well, to be honest… We've got a set of characteristics and criteria that Ofsted will look at to measure progress which probably won't correlate with this… There's a lot of stuff in this that is not included in Ofsted criteria… But it should be… But the stuff that is going well here… The culture of learning and the ethos will come out strongly… Even though we've got some things like year tens not doing so well…* [The year eight group when Ofsted failed the school December 2013 – previously discussed during year by year analysis above.]

173. **A1:** AH2 **R2/P2/T1: ACT:** {**Iy** Gently presented critique.} *I think it would be really interesting to cross reference some of the variables you have discussed… So I think it would be really interesting to get yourself* [AUT] *or Sean* [CT1] *… Or whoever it was… Looking at the things you've talked about… In a lesson… But have other people in the room who are experts in the other stuff like the school's adviser Mick Waters and maybe the head of department looking at the content of what's being delivered… To then like say there's three things there… Is this really having an impact on the progress… Or is it this… Or is it that… And how much is it like… Do you understand what it is I am trying to say?*

174. **A1:** AUT **T1: ACT:** {**Py** Partial agreement.} *I do… I do…* [You AH2 are sceptical of the variables chosen for this study, as per line 76 also.] *There's an evident caveat to this study… One of the things this study is doing… This way of working… Is to cut things down to a very small number of quantifiable variables… Like on-task behaviour… Taking that as my most important and dependent (so-called) variable… But if you look from the eighties… There have been many different ways of looking at classroom observation…*

175. **R1:** [Continues.] *You have O'Leary* [O'Leary, 2014] *at the University of Wolverhampton who is very sceptical of the expert-view of the classroom... Including Ofsted so-called experts… He is probably the most recent commentator… But from a pedagogic viewpoint … He's not a psychologist… To have published on the whys and wherefores of classroom observation…*

176. **R1:** [Continues.] *Probably the difference between the psychological and educational researcher… If they can usefully be distinguished… It seems to me to be the case… From my reading of the research… The psychologist is generally trying to simplify things down to a few significant variables whereas the educationalist appears to try and include as many variables… Measurable and immeasurable… In their studies.*

177. **A1:** AH2 **A2: ACT:** {**Py** Admission of scepticism.} *I'm playing devil's advocate… I'm not trying to be awkward…* [Yes, I am.]

178. **A1:** AUT **A2: ACT:** {**Py** Partial agreement.} *No… No…* [Of course, not!]

179. **A1:** AH2 **P2: ACT:** {**Iy** New competing hypothesis.} *You could be looking at a lot of passive learning…* [**A1:** HKS4 **A2: ACT:** *I was going to say that…] For example, you could see a lot of compliant behaviour… It could just be personality traits… You do get lots of students like that… Who are just naturally very quiet… You know… People… And who are on-task but what they're doing is nowhere near the level it needs to be… And that's why I think it would be really interesting if… Like… A psychologist was looking at it… Then you'd have the department leads… And the other link people saying yeh… And the content is good… And the challenge was good… We had a bit of that when Mick came in…*

180. **A1:** CT1 **P2/T1: ACT: {Iy** New competing method.} *You could take a more qualitative approach, I would think…*

181. **A1:** AUT **T1: ACT:** {**Py** Partial agreement. Then defence – lines 181 - 183.} *Yeh… Yeh… Yeh… You are absolutely right… That is an option… But what happened was that Ted Wragg and many educationalists before him… Developed complex observational schedules… I don't know whether anybody here remembers these… A huge number of tick-boxes to measure what was going on in the classroom… Designed to capture some of the rich stuff that was going on … The curriculum and the engagement… The pace of the lesson and so on… But… What happened is that such schedules became unwieldy… And unreliable… Test, re-test reliability was very poor…*

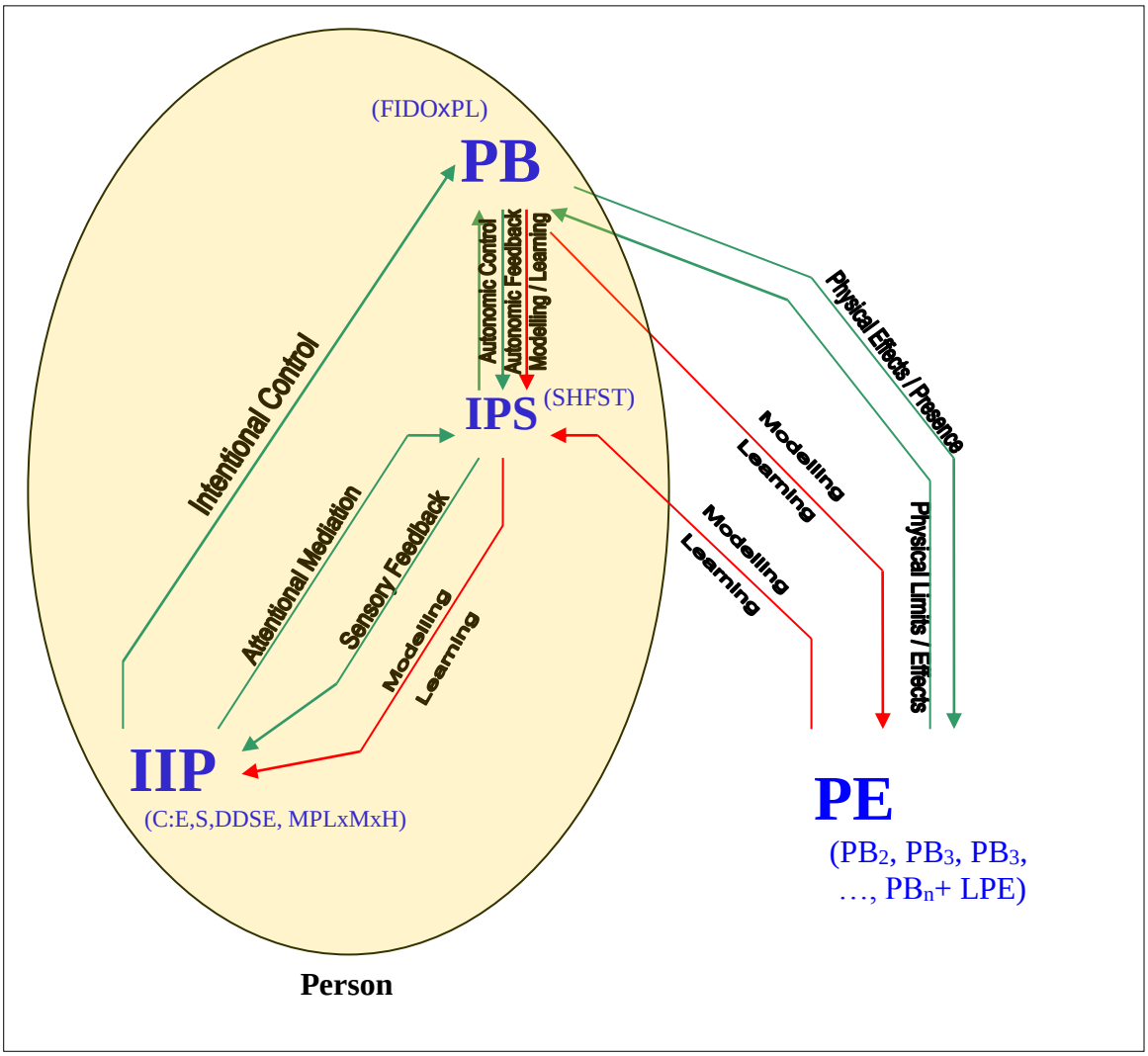
182. **R1:** [Continues.] *Politically, there was a reaction… Now Ofsted gives less structured guidance… To the extent that Ofsted inspectors use a very limited set of qualitative headings to try and capture all this stuff…. A wide range of variables in a quick visit to the classroom… But is this a robust or valid way of doing classroom observations… To try and capture so much almost by a quick feel… Using a few vaguely defined indicators… In a short visit to your classroom…*

183.**R1/E1:** [Continues.] *So instead… Here… I have gone back to a quantitative psychological approach… Measuring one or two valid variables in a robust and reliable way…* [Pause.] *This is a limited simple approach… But I know more precisely what I am trying to measure…*[Pause.] *I know we have got to finish now… It's late…*

184. **A1:** HT **R2: ACT:** *Yes. We've got to finish now… But this is food for thought. This is a different way of doing things and it has shown us different aspects… Thank you so much for coming today… Whether we can do much with this before Ofsted's visit, time will tell… But I definitely want to address the way… And how much… We talk to students… We do need to get better at this…* [**A1:** ALL **A2:** [Murmurs of agreement.]] *And we'd like to have you back… Maybe to help… Support Sean* [CT1]*… Re-run this with the maths department… To show how much they have moved on…* [**A1:** CT1 **A2: ACT:** *Mm… Yes…*]

185. **A1:** ALL **A2:** [Thanks, mutual praise statements and departures.]

[End of transcript.]

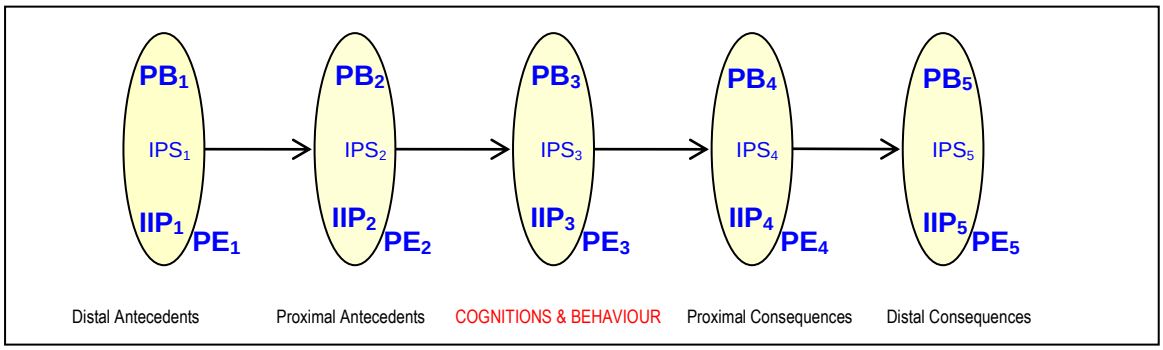


New Reciprocal Determinism

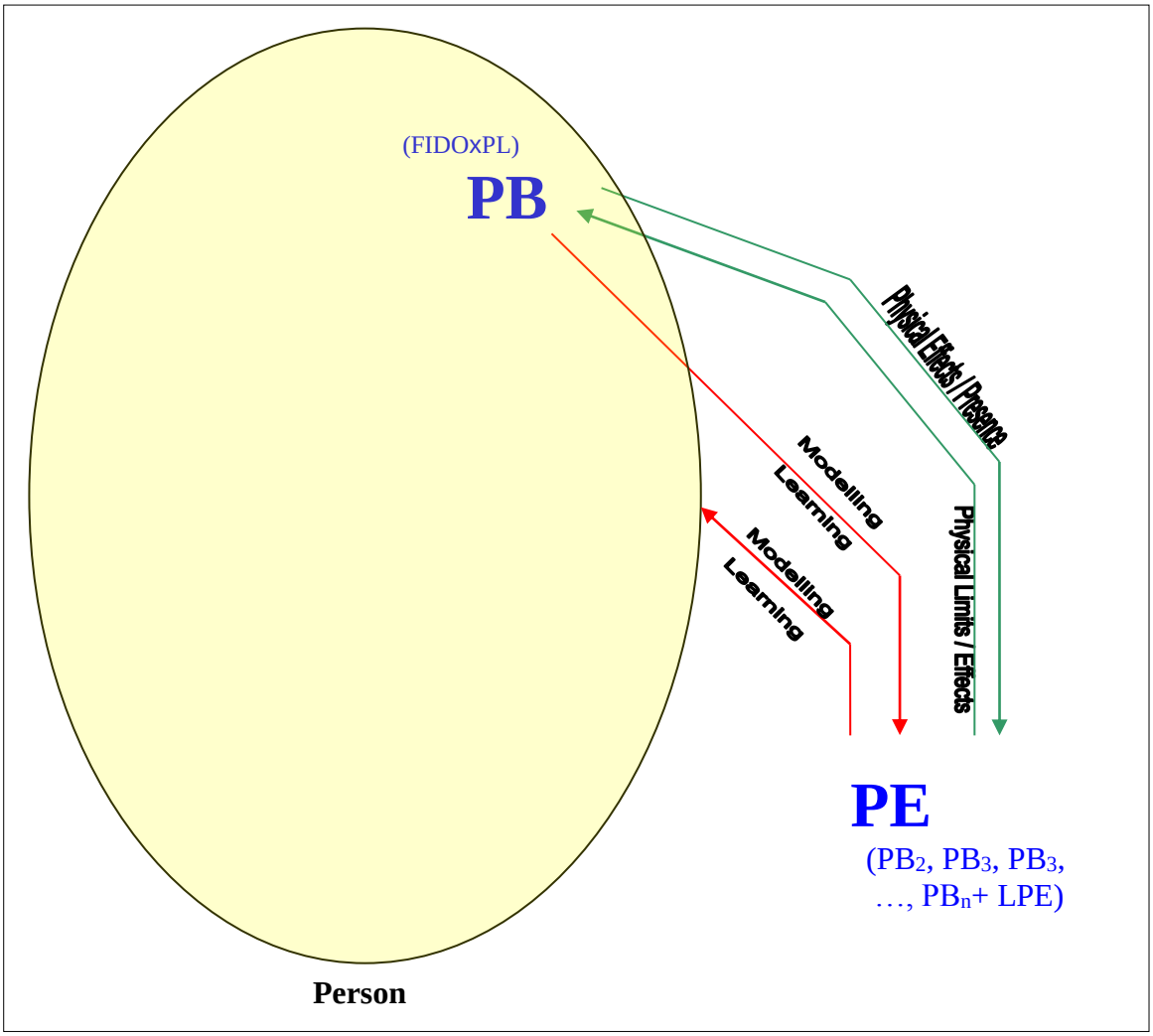
(Apter, 2007, Bandura 1977)

NOTES:

1. **PB** = **P**erceptible **B**ehaviour = **F**requency/**I**ntensity/**D**uration/**O**nset[s] inc. cycles x **P**hysical **L**imitations including appearance



New Reciprocal Determinism over time



Perceptible Phenomena

NOTES:

5. **PB** = **P**erceptible **B**ehaviour = **F**requency/**I**ntensity/**D**uration/**O**nset[s] inc. cycles x **P**hysical **L**imitations including appearance
6. **PE**=**P**erceptible**E**nvironment = **PB$_2$, PB$_3$, …, PB$_n$**= **P**erceptible **B**ehaviour of other people + **LPE**=**L**imitations of the **P**hysical **E**nvironment in which **PB$_1$** is

SURVEY OF UK SECONDARY CLASSROOM (Y7-Y11) BEHAVIOUR - SUKSCB

OVERVIEW

Thank you for taking part in what will be the largest observational survey of secondary school student and Teacher classroom behaviour attempted to date. It follows a similar study of primary school behaviour - the 'Mass Observation of Primary Classroom Behaviour', (Apter, Arnold and Swinson, 2010, *Educational Psychology in Practice,* 26, 2, pages 151 to 171). The findings of that Primary phase study were widely reported in national newspapers, and by *DEMOS,* the cross-party policy and research think-tank (http://www.demos.co.uk/) and by the BBC. The research design of the current Secondary phase study has been examined and vouchsafed by the University of Cardiff's School of Psychology's Ethics Committee.

With the current study, I am interested in investigating a number of research questions including:

- Do previous findings about ratios of social and academic praise setting currently apply in the secondary school setting?
- Are there regional variations in data?
- Are there 'time of day' variations in data?
- Are secondary phase student 'on-task' rates and classroom behaviour better or worse than previously reported?

SUKSCB 4 STEPS

1. Obtain materials by emailing: SUKSCB@aol.com Please do not send results to this address. It is for project enquiries and registration only;
2. Preparation for classroom observation (Teacher Selection, Teacher Preparation, Student Selection);
3. Complete 2 x observations of 20-30 minutes each, one a.m. and one p.m. ;
4. Transfer observation results to the SUKSCB Data Summary Sheet and email the results to: SUKSCBresults@aol.com .

The period for you to complete your observations and email them back to the SUKSCB Project has been extended and is between now and **20.7.2014**. On receipt, your name will be placed in a weekly draw for a £10 gift/book token. These tokens have been donated as incentives for SUKSCB observers by the University of Cardiff.

If at any time you require any clarification of these instructions or have any further questions please e-mail SUKSCB@aol.com.

Thank you for taking part.

Brian Apter

SUKSCB Project

GUIDANCE NOTES

STEP 1 Materials and familiarisation

In order take part in this survey and to carry out the observation you will need to send an email expressing your interest to SUKSCB@aol.com . The 4 documents that you require will be sent to you by return. [You can also download the documents from the EPNET file cache labelled SUKSCB.] The 4 documents are:

A. SUKSCB Overview and Guidance Notes
B. SUKSCB Letter to Head Teachers (exemplar template)
C. *MICRO* Recording Sheet
D. SUKSCB Data Summary Sheet

You will need to familiarise yourself with all the materials - especially the *MICRO* observation recording sheet and the instructions for use (files A and C). Incidentally, *MICRO* stands for Mixed Interval Class Room Observation. If you have not used a classroom observation schedule like this before, it is helpful to have a practice run before attempting to use it for this research.

This will be an ethically and methodologically rigorous study and it is important that you follow the procedures described below.

STEP 2 Preparation for observations

2.1 Teacher selection

The school chosen will probably be a familiar mainstream school within a group of schools that you regularly work in. You will need both a Head Teacher and the Teacher(s) you intend to observe to agree to the school's involvement in SUKSCB. The SUKSCB Letter to Head Teachers (download document B) is a suggested letter for Head Teachers which explains the project and aspects of confidentiality and ethics. You may alter letter-headings and layout to suit your professional style or the style-sheet of your EP service but the content should fundamentally be the same.

- Choose standard academic lessons where the students are mostly seated such as with English, maths or science. Do not choose lessons with large proportions of student-movement or equipment-dependent activities such as: P.E., music or I.T.
- Avoid choosing a school (or classes) where there are exceptional behaviour or class-control issues or a school which has been placed in 'Special Measures' by Ofsted.
- A 20-30 minute observation of a teacher in the morning and a 20-30 minute observation of the same teacher during the afternoon session would be *ideal*, but it is *not essential* that the same teacher is used for both observations.
- You may decide with the Head Teacher to complete more than one pair of observations in the school.

2.2 Teacher preparation

Once the teacher(s) and classes have been selected you need to:

- Reassure the teacher(s) that they have not been chosen for a critical purpose.
- Explain that all recorded aspects of the observation are anonymously and confidentially stored and will only be used in a generalised unattributed form for the SUKSCB project.
- Emphasise that teacher(s) have the right to not take part and that they have the right at any time during the observation to ask you (the observer) to curtail the observation and leave the classroom.

- Explain that teacher(s) have a right to see the record of observations at the end of the day.
- It is important that you do not discuss with teacher(s) the detail of your observations prior to completing all the observations in a particular school. If teachers have advanced knowledge that you are recording their verbal behaviour then they may well alter their behaviour. It is sufficient to use the following form of words: *'The observation will be concentrating on student behaviour and the classroom variables that influence behaviour.'*

2.3 Student selection

Use 5 students in the morning session, selected randomly; and 5 in the afternoon, selected randomly.

- It does not matter if the observed students are the same in the morning and afternoon sessions but it is important that students are chosen at random for each observation.
- One quick way observers have used to choose random selections of 5 students is by privately assigning numerical identifiers to a list of all students in a class, e.g. 1 to 25, and then asking a teacher in the school staffroom who is not associated with the observation task to pick 5 numbers at random, between 1 and 25.
- The teacher you are observing should be unaware of which students you will be observing as this may result in the teacher giving that student more attention than would otherwise be the case.

STEP 3 Observations

The following is a set of instructions for using the *MICRO* recording sheet.

3.1 Fill in details at top of form - **Students Initials**, **Class / Teacher** identifier, **Date**, **Type of Lesson** (circle type), **Learning Support Assistants' ID(s)** (number of, or initials) and **Class / gp. size**. Enter '**START TIME**' as 24 hour clock time (hh:mm e.g. 09:35).

NOTE: Individual identifiers are only suggested for your convenience. The survey protocol requires that all staff and students are anonymous to the researcher. The completed *MICRO* is not required by the researcher. Data is returned to the researcher on the anonymous 'SUKSCB Data Summary Sheet' (see below.)

3.2 Locate in the classroom your randomly selected target students - **S1**, **S2**, **S3**, **S4** and **S5**. The **S6** column is for your observations of the main class teacher (or adult Learning Support Assistants if they take over the main teaching role.).

3.3 Decide on the time frequency of your observations. Try 2 minutes for each row to begin with. With increased experience of using the *MICRO* you will be able to complete rows of observations every minute. Using your wrist watch or a stopwatch, you should enter the minutes: 0, 2 (e.g. start time and start time plus 2 minutes) and then 4, and 6, and so on, down the **+T** column as you observe the lesson. The *MICRO* form will cover an hour of 2-minute observations (30 rows of 2 minutes each), or half an hour of 1-minute observations.

3.4 Now quickly look at your first student **S1**. Is s/he **f**ollowing the last **d**irection (**fd**) given him or to the class by the teacher as far as you able to observe? Mark a tick (✓) or cross (**x**) in the **fd** column according to your best judgement of their behaviour.

OPTIONAL: When not involved in completing a *MICRO* for the SUKSCB project, you might wish to record what the first student **S1** is *actually* doing each time you observe them. There is space on the *MICRO* observation form for this purpose. This information is not required for SUKSCB but might be useful to you if you have a particular interest in the specific student's (**S1**) behaviour when using *MICRO* for a different purpose with the

teacher – perhaps when asked to observe a behaviourally challenging student. For example: you might fill-in a code such as '**QW**' when **S1** is **Q**uiet **W**orking or '**TP**', when they **T**alking to **P**eer, in the 'code' **C** column. Please see the first column of *MICRO* form for suggested codes. You will see that there is additional space on the *MICRO* form to create new codes that you can invent as you require.

3.5 Q*uickly* repeat process for target students **S2**, **S3**, **S4** and **S5** in respective **fd** columns.

NOTE: There is no option for recording a behaviour-code for **S2**, **S3**, **S4** and **S5's** behaviour, in contrast to student **S1**.

3.5 For remainder of the observational time-slice, you will observe the class teacher (or Learning Support Assistant if they have taken over the main *teaching* role during your observation) and record their verbal behaviour in the 5 columns under the **S6** heading.

FOR EXAMPLE: if you are using a 2-minute time-slice, you might have used 30 seconds to observe and record **fd** (or not) for **S1**, **S2**, **S3**, **S4** and **S5**. This would leave 1 minute 30 seconds to observe and record the teacher **S6**.

3.7 The column, **INX**, will give an indication of how much whole-class neutrally-toned teacher-talk there is during an observation. It is a percentage estimate of the amount of neutrally-toned verbal behaviour (neither critical nor praising) used by the teacher during the time-slice. To aid your estimate, you can make up to 5 tally-marks - approximately one per each 15 to 20 seconds - in the **INX** column during the process of observing the teacher using **IN**structions, e**X**planations and/or e**X**positions (**INX**) to the class, to groups of students or to individual students.

NOTE: 5 is the *maximum* count in an **INX** box. If you reach 5 in any **INX** box then I will assume that you have observed continuous teacher-talk occurring during the remainder of that time-slice. This makes it possible to estimate a % figure for **INX**, i.e. **INX%** (where each tally mark is assumed to be equal to 20% of the time-slice).

OPTIONAL: You might wish to annotate the record with your own comments for your own purposes – perhaps to help you give feedback to the teacher later. The **OBS NOTES** column to the far right of the *MICRO* sheet provides you with a small amount of space for your comments. The SUKSCB project does not require you to fill in this column and there is no space for notes on the SUKSCB Data Summary sheet.

3.8 Simultaneously but separately, tally the teacher's use of praising or critical comments. These might occur within their continuous teacher-talk or as discrete instances in otherwise silent periods when the teacher is not talking.

Tally individual instances of positive or negative teacher verbal behaviour by making marks in the columns: Task Performance Positive (**TPP**), Social / Behavioural Positive (**SBP**), Task Performance Criticism or redirection (**TPC**), or Social / Behavioural Criticism or re-direction (**SBC**).

FOR EXAMPLE: if a teacher is conducting an e**X**planation for the entirety of a time slot, but 'mid-flow' tells 'John, don't lean back on your chair!' there would be a tally of 5 (100% teacher-talk) in the **INX** box and a tally of 1 (one instance of social-behavioural criticism) in the **SBC** box.

NOTE: Whereas the **INX** *maximum* is always 5, simple tallies in the **TPP**, **SBP**, **TPC**, **SBC** boxes can exceed 5.

3.9 At the end of the time segment, begin the whole process again by observing the student **S1** again and recording your observations on the next row down, repeating steps 3.3 to 3.8 (above). Continue to repeat this cyclical process until the end of the observation period (minimum 30 minutes).

3.10 Examples of categorisation of teacher' comments

General note: Categorisation of student and teacher behaviour is not an exact science. Observers will rely on their best judgement of how to categorise a particular behaviour. For example, the observer must decide whether a student staring into space is pausing for thought and following directions *or* vacantly day-dreaming and not following directions. A statistically acceptable degree of inter-observer unreliability has been allowed and tested for in a reliability pilot-study during the design and development of the *MICRO* observational protocol.

3.11 **INX:** "Get your books out and find page nineteen", "Line up by the door", "Find and underline all the verbs", "This is the way to set out these calculations", " and "Here is a diagram of an internal combustion engine," are **IN**structions, e**X**planations or e**X**positions. Typically, **INX** are *neutrally* delivered instructions, academic commentaries or descriptions being given for the first time to groups of students at the beginning or during academic activities or at activity-transition points by a class teacher (or a substituting TA).

3.12 **TPP:** "Well done, class, good work", "Darren, that's a really good argument you have made about Shylock's motivation", and "Sophie, that's lovely writing", are all examples of **T**ask **P**erformance **P**ositive comments. Typically, they are enthusiastic or positive recognition / praising comments addressed to students about outcomes from a specified activity that has been directed, organised or sanctioned by the class teacher.

3.13 **SBP:** "Good, 7N. It was great to hear a lot less chat today and see a lot more work happening", "Thanks Darren for *helping* Paul with his spellings", and "You are cooperating as a research group and working well together", are all examples of **S**ocial **B**ehavioural **P**ositive comments. Typically, they are enthusiastic or positive recognition / praising comments to students in respect of their pro-social behaviours or compliance with instructions or rules that an adult has given them.

3.14 **TPC:** "John, I am a bit disappointed that you haven't completed that worksheet yet", "You need to up your work rate, if you are going to be ready for the exam next week", and "Sophie, remember. The examiner needs to be able to read your writing" are all examples of **T**ask **P**erformance **C**riticism.

NOTE: **TPC** includes *repeated* directions (redirections). Typically, they are *implicitly* critical or corrective comments to students about outcomes from a specified academic activity that has been directed, organised or sanctioned by the class teacher.

3.15 **SBC:** "Stop leaning back on your chair, John and interfering with Phillip. He's trying to work, even if you are not", "I need to remind you Blue Table that the rule is 'keep your hands and feet to yourself'", and "Kelsey. Enough! I won't tolerate swearing", are all examples of **S**ocial **B**ehavioural **C**riticism or re-direction comments. Typically, they are *corrective* comments and *repeated* directions addressed to students about anti-social, non-compliant or unacceptable behaviours by an adult.

3.20 Easy Step by Step Calculation of Averages

3.21 At the bottom of the *MICRO* recording sheet there are 2 grey rows with 13 vacant white boxes to fill in. This is for the anonymous data required by the researcher. Calculate the time as a percentage that the observed students were 'On-Task'. To do this, add **S1(fd)** column ticks, divide by the number of observations and then multiply by 100. Put this number **fd%1** into the labelled box at the bottom of the *MICRO* form.

3.22 Repeat this procedure for **S2(fd)**, **S3(fd)**, **S4(fd)**, **S5(fd)** columns. You now have 5 'following directions' percentages: '**fd%1**', '**fd%2**', '**fd%3**', '**fd%4**' and '**fd%5**'n the boxes at the bottom of the recording sheet.

3.23 To obtain average student 'following directions' time-estimate for the class that you observed as a percentage, add **fd1%** + **'fd2%** + **fd3%** + **'fd4%** + **fd5%** and divide by 5.

3.24 Transfer this figure into the On-Task (**ON%**) box of the SUKSCB Data Summary Sheet. Note: If it is a morning observation, you will fill in **Section 2**, left-hand column **a.m.** of the SUKSCB Data Summary Sheet or right-hand column **p.m.** for afternoon observation data.To obtain the average student 'Off-Task' time as a percentage, subtract the same figure (**ON%**) from 100. Transfer this figure into the Off-Task (**OFF%**) box on the SUKSCB Data Summary Sheet.

3.25 Count up: teacher's **IN**structions, e**X**planations or e**X**positions **INX**. There is an imposed maximum of 5 per time slot so it is possible to calculate a percentage teacher talk time **INX%**.

FOR EXAMPLE: For a 32 minute observation there are 16 x 2 minute time slots. Pupils are observed for 30 seconds in each time slice, so their teacher is observed for 1 minute 30 seconds in each time slice. The maximum number of tally marks would still be 16 x 5 = 80. A total of 60 tally marks in the **INX** column would be the equivalent of 75% (60/80): the teacher would have been estimated to have been speaking for 75% of the observed time period.

3.26 Count up: Task Performance Praise comments from **TPP** column, Social Behavioural Praise comments from **SBP** column, Task Performance Criticism comments from **TPC** column, and Social / Behavioural Criticism from **SBC** column. Enter totals into boxes at the bottom of the *MICRO.*

3.27 Calculate ratios: Task Performance Praise comments to Social / Behavioural Praise (**TPP:SBC**); and Praise to Criticism ratio (**TPP+SBP**):(**TPC+SBC**). Enter into boxes, bottom right-hand corner of *MICRO.*

STEP 4 Transfer Observation Data to SUKSCB Data Summary Sheet and Email or Post

4.1 Record details of the **School Context** for the 2 observations (a.m. and p.m.) on the *SUKSCB Data Summary Sheet*:

- Observer name
- Educational Psychology Service
- Post code of School
- Locale type, e.g. Inner City OR Town/Suburban OR Rural/Village
- Free School Meals (as percentage of school roll)
- School size as number on school roll

There are then two columns for **Class Context** information, morning and afternoon:

- Date of observation (dd / mm / yy)
- Year group (Yr 7 – Yr 11)
- Gender of teacher
- Total years teaching
- Time of day of observation, 24 hour clock (hh / mm)
- English / Maths / Science / Other
- Number of Adults in room (not including observer)
- Number of Students in class

4.2 Transfer **INX%**, **TPP**, **SBP**, **TPC** and **SBC** into the appropriate boxes on the SUKSCB Data Summary Sheet.

4.3 When you have finished a pair (a.m. and p.m.) of planned observations, you should have in your possession *one* completed SUKSCB Data Summary Sheet in respect of each *pair* of observations completed.

You can, of course, email or post the data from as many pairs of observations as you are able to complete on separate SUKSCB Data Summary Sheets.

Please email electronically completed SUKSCB Data Summary Sheets to SUKSCBresults@aol.com or post paper copies to: **Brian Apter, District Educational Psychologist, Jennie Lee centre, Lichfield Road, Wednesfield, WV11 3HT**

4. Example of completed *MICRO* for a 30 minute observation:

MICRO Recording Sheet

Mixed Interval Class Room Observation

Student Initials (S1/S2/S3/S4/S5): JS/ST/AM/PL/TR Class / Teacher: SL Date: 13/4/08

Type of lesson: (Eng) / Maths / Sci. / Other LSA(s) ID(s): x2 IJ / BA Class / gp. size: 25

COL C CODES	+T	S1 fd	S1 C	S2 fd	S3 fd	S4 fd	S5 fd	S6 INX	S6 TPP	S6 SBP	S6 TPC	S6 SBC	START TIME: OBS NOTES
QW=Quiet Working	0	✓	LT	✓	✓	✓	✓	IIII					Exposition of lesson by teacher - went through worked model
OS=Out of seat	2	✓	LT	✓	X	✓	✓	IIII					
TP=Talking to peer	4	✓	LT	✓	X	✓	✓	IIII					
TT=Talking to teacher	6	✓	LT	✓	X	✓	✓	IIII	II				
OE=Occupied with equipment	8	X	TP	✓	X	X	✓	I		III		IIII	S2 asked teacher point of clarification
LT=Listening to teacher	10	X	TP	✓	✓	✓	✓						
WQ=Waiting in queue	12	X	TP	✓	X	✓	✓		I			IIII I	
HU=Hand up waiting in seat	14	X	TP	X	X	✓	✓						
GS= Gazing into space	16	X	TP	✓	X	✓	✓		II				
CL=On carpet listening	18	X	TP	✓	X	✓	✓		I			II	S1 did not appear to understand work
HD= Head on desk	20	X	OE	✓	X	✓	✓						
PB= Picking behaviours	22	✓	TT	✓	✓	X	X	II	I			III	
TA= Individually working with TA	24	✓	QW	✓	✓	✓	✓		I				
Etc.	26	✓	QW	✓	✓	✓	✓		II	I		II	S1 asks for clarification
	28	X	OE	✓	✓	X	✓		III				Then S1 begins to work
	30	✓	QW	✓	✓	✓	✓		II			I	S6 says "I've shown you already.... were you not listening?"
Totals	✓=S			15	7	13	15	23	15	4	0	19	
		8/16		15/16	7/16	13/16	15/16	23/80 (0%)					
NB	S1 was not the biggest problem in class whatsoever. Main findings not Teacher needs to increase SBP - feedback to teacher.												suggested?
TOTALS	Avg. On-task = sum(fd%1,...,5)/5=	50	+	96	44	81	94	28.75 INX%	15	4	0	19	Ratios: TPP/SBP =TPP:SBP 15:4; & P/C = (TPP+SBP):(TPC+SBC) 19:19
	70.25								Raw Tallies				

OTHER CODES: Time (T) = In minutes, following direction (fd): ✓ or ×, **INX** first time instruction, or explanation, or exposition; neither critical nor praising (Max: 5 per box.) **TPP** task performance praise **SBP** social/behavioural praise, **TPC** task performance criticism or redirection **SBC** social/behavioural criticism or redirection

6. Example of Completed SUKSCB Data Summary Sheet ready to be emailed or posted:

SUKSCB MICRO Data Summary Sheet

Survey of UK Secondary Classroom Behaviour

Section 1 – School Context

• Observer name	Brian Apter
• Educational Psychology Service	Hedgford
• Postcode of school	RV69 1UR
• Locale type, e.g. Inner City or Town/Suburban or Rural/Village	(IC) / TS / RV* delete
• Free School Meals (% school roll)	27 %
• School size as number on roll	940

Section 2 – Class Context

	a.m.	p.m.
• Date of observation (dd/mm/yy)	13/4/08	14/4/08
• Year group (Yr 7 – Yr 11)	7	9
• Gender of Teacher	(M) / F*	(M) / F* Same
• Total years teaching (NQT= 0)	14	14
• Time of day of observation, 24hr clock (hh:mm), e.g. 15:30	9.30	14.40
• English / Maths / Science / Other	E / (M) / S / O*	E / M / S / (O)*
• Number of adults in room	3	1
• Number of students in class	25	15

Section 3 – Observation Data

	a.m.	p.m.
• On-Task ON%	78.25 %	63 %
• Off-Task OFF%	21.75 %	37 %
• Percentage Teacher neutral verbal behaviour (INX%)	28.75 %	14 %
• Task Performance Praise comments (TPP) tally	15	4
• Social Behavioural Praise comments (SBP) tally	4	2
• Task Performance Criticism or redirection comments (TPC) tally	0	1
• Social Behavioural Criticism or redirection comments (SBC) tally	19	1
• Duration (minutes) of observation	30	25

Email electronically completed SUKSCB Data Summary Sheets to SUKSCBresults@aol.com or post paper copies to:
Brian Apter, District Educational Psychologist, Jennie Lee centre, Lichfield Road, Wednesfield, WV11 3HT

7. Example of completed letter to Head Teacher (observation request):

Survey of UK Secondary Classroom Behaviour

To: The Head Teacher,

Bogstandard Comprehensive,
Notwithstandington,
Rutland,
RU69 1UR

From: Brian Apter,

The Hedgford Bunker,
Hedgford,
Rutland,
RU69 1RU

Date: 1st April, 2011

Dear Head Teacher,

The Survey of UK Secondary Classroom Behaviour (SUKSCB) is a national project that is being conducted by a researcher from the University of Cardiff. The project is aiming to observe and analyse student behaviour in academic lessons in Secondary Schools across the United Kingdom.

We are looking in detail at the relationship between different lessons, the time of day, teaching style and student behaviour. This type of study has been conducted before in secondary schools but in the past has been restricted to relatively small geographical areas. With this national study we are hoping to detect if there are significant regional variations in the collated data.

All data sent to the researcher will be anonymously stored. Individual teachers, schools and students cannot be identified by the researcher. The research design has been examined and passed by the University of Cardiff's School of Psychology's Ethics Committee.

As this is a whole-class observation project and carried out in the strictest confidence, we are advised that it is not necessary to seek parental permission. However, at your discretion, you may wish to let parents know in general terms that the school is taking part in a nationwide survey.

Participation in the project is voluntary. Permission for any observations in your school must be at your discretion and of course the class teachers themselves must agree to the observation.

The two observations take half an hour to complete. We have asked the psychologists to observe, where possible, the same teacher and class in a morning and afternoon session. We are keen to have a valid sample and therefore the class/classes to be observed should ideally be selected at random.

We intend to publish the results of the survey but no schools will be named. We will let you know when that happens and what we have found out.

If there is any further information that you require the University of Cardiff researcher can be contacted by email by your psychologist.

We hope that you will be able to agree to your school participating in the project.

Yours sincerely,

Brian Apter
Hedgford Educational Psychology Service