ELSEVIER

Learning and Instruction 21 (2011) 715–730

Learning and Instruction

www.elsevier.com/locate/learninstruc

Examining the effect of class size on classroom engagement and teacher–pupil interaction: Differences in relation to pupil prior attainment and primary vs. secondary schools

Peter Blatchford*, Paul Bassett, Penelope Brown

Department of Psychology and Human Development, Institute of Education, University of London, 25 Woburn Square, London WC1H OAA, UK

Received 20 September 2010; revised 15 March 2011; accepted 1 April 2011

Abstract

It is widely recognized that we need to know more about effects of class size on classroom interactions and pupil behavior. This paper extends research by comparing effects on pupil classroom engagement and teacher–pupil interaction, and examining if effects vary by pupil attainment level and between primary and secondary schools. Systematic observations were carried out on 686 pupils in 49 schools. Multilevel regression methods were used to examine relationships between class size and observation measures, controlling for potentially confounding factors like pupil attainment. At primary and secondary levels smaller classes led to pupils receiving more individual attention from teachers, and having more active interactions with them. Classroom engagement decreased in larger classes, but, contrary to expectation, this was particularly marked for lower attaining pupils at secondary level. Low attaining pupils can therefore benefit from smaller classes at secondary level in terms of more individual attention and facilitating engagement in learning.

Keywords: Class size; Classroom processes; Teacher pupil interaction; Classroom engagement; Classroom observations

1. Introduction

In many countries over the world there has been a widely reported debate over the educational consequences of class size differences. Opinions vary from those academics and policy makers who argue that class size reduction is not cost effective to those who argue that it should be a main feature of educational policy. In some countries policy has changed in favor of small classes. In the U.S.A, over 30 states have enacted legislation for class size reduction (CSR) programs. Current Government policy in England and Wales is for a maximum class size of 30 for pupils aged 4–7 years, and larger cuts are planned in Scotland. There have been initiatives involving class size or pupil to adult ratio reductions in the Netherlands and New Zealand. In East Asia, many countries and cities (including Shanghai in the Chinese Mainland, Hong Kong, Macau, Taiwan, Korea and Japan) have implemented 'small class teaching' initiatives.

Most attention has been paid to whether or not smaller classes lead to better academic outcomes for pupils. There is a good deal of controversy over the magnitude of these effects (see reviews by Anderson, 2000; Biddle & Berliner, 2002; Blatchford & Mortimore, 1994; Blatchford, Goldstein, & Mortimore, 1998); Blatchford, Russell and Brown, 2009; Ehrenberg, Brewer, Gamoran, & Willms, 2001; Finn, Pannozzo, & Achilles, 2003; Galton, 1998; Grissmer, 1999; Hattie, 2005 and Wilson, 2006), though there is some agreement, drawing on experimental (e.g. Finn & Achilles, 1999) and naturalistic studies (Blatchford, Bassett, Goldstein, & Martin, 2003), that smaller classes have positive effects on pupil academic performance, if introduced immediately after school entry, that is, with the youngest children in school. However, it is now recognized by many - and not just critics of class size reductions - that in order to better understand the effects of class size, and help facilitate better

* Corresponding author. Tel.: +44 (0)20 7612 6268; fax: +44 (0)20 7612 6304.

E-mail address: P.Blatchford@ioe.ac.uk (P. Blatchford).

0959-4752/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.learninstruc.2011.04.001

classroom environments and effectiveness, we need to know more about effects on what goes on in classrooms, that is, classroom 'processes' such as interactions between teachers and pupils and pupil behavior. This is of importance in its right and the main aim of this paper was to research in a systematic way the relationships between class size and classroom processes.

This paper examines effects of class size on two key processes: pupil classroom engagement and teacher–pupil interactions. It extends previous research in five ways: 1. it contrasts effects on the two processes, i.e., pupil classroom engagement vs. teacher–pupil interactions; 2. it examines whether effects vary by pupil prior attainment; 3. it examines whether effects vary in primary vs secondary schools; 4. it addresses whether there are class size thresholds above and below which effects are more marked; and 5. it uses a rigorous systematic observation method for recording ongoing classroom behavior.

1.1. Effects on teaching vs. classroom engagement

Effects of class size on classroom processes tend to fall into two main camps. First, there are those concerned with teacher to pupil interactions. It seems likely that bigger classes will decrease the amount of time that can be spent on instruction and dealing with individual child. This is consistent with teachers' views (Bennett, 1996; Pate-Bain, Achilles, Boyd-Zaharias, & McKenna, 1992), and some previous research (Achilles, 1999; Anderson, 2000; Blatchford, Moriarty, Edmonds, & Martin, 2002; Bruhwiler & Blatchford, 2011; Cooper, 1989; Glass, Cahen, Smith, & Filby, 1982; Molnar et al., 1999). However, observational research by Shapson, Wright, Eason, and Fitzgerald (1980), and reviews by Ehrenberg et al. (2001) and Slavin (1989) support the view that the effects of class size reductions on teaching are minimal.

The second set of factors related to class size differences concerns pupils' classroom engagement. Finn et al. (2003) argue that the effects of class size on pupils' classroom engagement are more important than those on teaching. In the English CSPAR research, a connection with pupil 'on task' behavior was found with the younger pupils (aged 4/5 yr olds) but not older primary pupils (aged 10/11 years) (Blatchford, Bassett, & Brown, 2005).

There is then dispute about the balance of effects on teacher pupil interactions and pupil classroom engagement and this paper compares in a systematic way the effects of class size on each. More information on the specific aspects of classroom engagement and teacher to pupil interactions studied is given later in this introduction, after the other key features of the study are described.

1.2. Pupil prior attainment

Research has found that effects of class size on academic outcomes can vary by pupil characteristics such as prior attainment level, disadvantage and minority group status (Blatchford et al., 2003; Ehrenberg et al., 2001; Finn & Achilles, 1999; Molnar et al., 1999). Research in England found that the initially lowest attaining pupils benefited most from small classes in the first year of school (Blatchford et al., 2003). However, in a re-analysis of the STAR data, Konstantopoulos (2008) found that it was higher ability students who benefited most from small classes and small classes did not reduce the achievement gap. One of the justifications of small classes is the hope that it will help those with most ground to make up academically receive more individual attention and help them concentrate. Conversely a problem with large classes might be the adverse effect on the levels of classroom engagement of low attaining pupils. In this study we therefore wanted to see whether effects on classroom interactions and pupil engagement varied by initial attainment level.

1.3. Age effects

Research suggests that the youngest pupils benefit most in terms of academic outcomes from small classes (Blatchford et al., 2003; Ehrenberg et al., 2001; Finn & Achilles, 1999), e.g., because they better help children adjust to school and receive individual attention. But most evidence comes from primary aged pupils, and we know very little about effects of class size on older school pupils, i.e., secondary age pupils aged 11–16 years. Still less is known about age differences in class size effects on classroom processes such as teacher and pupil interactions and classroom engagement. There is a suggestion that effects become weaker with age (Finn et al., 2003) but this also requires further study involving older pupils. One of the features of this paper is that it examines, within one study, the effects of class size on classroom processes at four age levels that cover the whole of the primary and secondary school years.

1.4. Class size thresholds

It is often assumed that class sizes need to fall below a certain number (the figure of 20 is often mentioned) before they can have an impact on educational outcomes. However, this is likely to have much to do with the class sizes chosen in research. The best know study of class size effects - the Tennessee STAR project - for example, pre-selected and compared classes of about 17 with class sizes of about 23, and this is probably a main reason why the mid-point between the two is seen as important. However, this range of class sizes is not common in many countries, even in the USA, and an alternative approach is to examine effects of class size across the full range of class sizes, rather than presuppose class sizes likely to be important. Very little is known about class size thresholds below or above which effects on classroom processes are evident. Shapson et al. (1980) compared a number of different class sizes but when the range of class sizes is wide, as it is in the UK, this kind of design can become unwieldy. It is therefore potentially more insightful and valid to employ naturalistic studies within which class sizes vary as in the real world of education, and which therefore allow estimates of effects across the full class size distribution.

1.5. Research methods

A theme to emerge from the reviews of research cited above is that there are methodological weaknesses in much

research in this area. Studies have been relatively anecdotal, with a reliance on teacher opinion and comment. Finn et al. (2003) have pointed out that there is relatively little strong systematic observational research which would permit objective study of relationships between class size, teaching practices and student behavior. They identify the observational study by Blatchford (2003b) as one of the few studies of sufficient quality to be included in their review. The current paper builds on this earlier work. In contrast to other forms of data collection it builds up data on the basis of careful recording of ongoing behavior (rather than, say, ratings or judgments). Criticisms of systematic observation have usually centered on validity issues (e.g. Delamont & Hamilton, 1986), but it can be a useful research tool when answering specific research questions where data are needed on relatively easily observed behaviors (Croll, 1986; McIntyre & Macleod, 1986; Yodder & Symons, 2010). This was the method used by Shapson et al. (1980) and by Bourke (1986) and, though both studies are widely cited, they are rather dated now.

A feature of this paper is that it also employs more sophisticated statistical analyses than many previous studies and examines effects of class size on moment by moment behaviors, while controlling for other possibly confounding factors, particularly student prior attainment. It is also important to recognize that observations are not always independent of each other and that multilevel statistical models are required that deal with the clustered nature of observation data within pupils and within classrooms within schools. More details are given in the method section below.

1.6. Contextual approach

There is an underlying assumption, in many studies of teaching, of a direct model, with teaching affecting pupils' achievements and learning in a causal way (Blatchford, Kutnick, Baines & Galton, 2003). But teachers do not meet pupils out of context, and class size, or the number of children in the classroom, can be seen as one contextual influence on classroom life, which plays a part in affecting the behavior of teachers and pupils. Class size is positioned as one such 'context' factor in Dunkin and Biddle's (1974) early seminal model of effects on teaching. The conceptual roots of this view can also be found in Bronfenbrenner (1979) and ecological psychology. The basic idea is that the classroom context has distinctive forces or 'signals', different to other contexts, which pull events and participants along with them (Kounin & Gump, 1974). Different aspects of the class environment, such as the number of pupils in a class, can exert distinct pressures on teachers and pupils.

1.7. This study

Two aspects of teacher and pupil behavior are considered in this study: pupil classroom engagement and teacher–pupil interaction. These, and the research predictions, are now explained in more detail.

1.7.1. Pupil classroom engagement and off-task behavior

One main element of children's successful adjustment to school is likely to involve their productive engagement in class, as reflected in the extent of their work related interactions with teachers, other pupils and when working on their own. Many studies support the view that engagement, active learning time, time on task or some equivalent term (e.g. Creemers, 1994) are key aspects necessary for educational success.

Common sense and logic suggest that with more children in the class there will be more potential for distraction, and more possibility of being off task. Conversely in small classes there will be more opportunities to engage children and keep them on task. Finn et al. (2003) developed a theoretical and empirical case for why student classroom engagement is the key process that explains why smaller classes lead to better attainment. They conclude that students in small classes in the elementary grades are more engaged in learning behaviors, and display less disruptive behavior than do students in larger classes. Cooper (1989) reported studies that support this view, but there are limitations in much of this research because of the often informal designs and reliance on teacher reports (Finn et al., 2003).

A different view comes from Bourke (1986) who in an Australian study did not find a class size effect on primary school student engagement. As we have seen, the CSPAR study found an age effect in the sense that 4/5 year old pupils showed more off task behavior in larger classes but no effects were found on pupil attentiveness in 10/11 year old pupils.

There are limitations in many conceptualizations of on and off task behavior, in that the generic terms may conflate separate forms of behavior. There are three main contexts through which pupil on and off task behaviors will be expressed: when with teachers, other pupils or when on their own. It is helpful to know if these are connected in similar or different ways to class size differences. In the earlier CSPAR study, in the case of 4/5 year old pupils, there was more off task behavior in larger classes, but especially more passive off task behavior - more disengagement - when working on their own (Blatchford et al., 2005).

On the basis of these results it was predicted that there would be more general off task behavior in large classes. There were not strong grounds for predictions about age or type of pupil effects, but it was anticipated that effects would be most marked at primary level, and be most evident for the lower attaining pupils.

1.7.2. Teacher–pupil interactions

There were four forms of teacher–pupil interaction examined in this study: 1. overall amount of teaching; 2. teacher individual attention to pupils; 3. pupil active involvement with the teacher; and 4. classroom control and management.

1.7.2.1. Overall amount of teaching. It seems likely that the number of children in a class will decrease the amount of time that can be spent on instruction and time spent dealing with individual children. This expectation is consistent with teachers' views (Bennett, 1996; Pate-Bain et al., 1992) and some previous research (Cooper, 1989; Glass et al., 1982).

However, Shapson et al. (1980) found that there were no statistically significant differences between class sizes for most teacher activities, and teachers did not alter the proportion of time spent interacting with the whole class, with groups or with individuals. This conflicted with teachers' own experiences and there was, therefore, a gap between professional judgment and observational research findings.

In the CSPAR research (Blatchford, 2003a; Blatchford et al., 2005) it was found that more teaching took place overall in smaller classes but this was restricted to the youngest age group studied (4/5 years). In the present study the aim was to clarify any possible longer term age effects by examining effects through primary and secondary school stages, and to see whether low attaining pupils, as might be expected, experience more teaching overall in smaller classes, in order to help them catch up.

1.7.2.2. Teacher individual attention to pupils. Perhaps the most consistent finding concerning class sizes effects on classroom processes is that reduced class size is related to individualization of teaching (Betts & Shkolnik, 1999; Ehrenberg et al., 2001; Molnar et al., 1999). Results from the CSPAR systematic observation studies (Blatchford, 2003b; Blatchford et al., 2005) showed that although there was a heavy reliance on whole class teaching and individual work in primary schools, pupils in small classes were more likely to experience one-to-one teaching and were more often the focus of a teacher's attention. In the same vein a multi-method study by the NICHD Early Child Care Research Network (2004) summarizes data by arguing that smaller first grade classrooms appear more child-centered than larger classes, and that larger classes become more structured. Anderson (2000) argued that small classes encourage a more personalized and appropriate curriculum for individual pupils.

A number of studies in the USA (Finn & Achilles, 1999; Molnar et al., 1999) and the UK (Blatchford et al., 2003) suggest that CSR tends to benefit lower attaining and disadvantaged pupils, and it might be expected that more individual support in smaller classes will be targeted at the lowest attaining pupils. However, it might be expected that as pupils progress through primary into secondary school, the more structured and centralized curriculum, and the preparation for public examinations, will reduce any effect of smaller classes on individualized attention. This study therefore examined whether effects of class size on individual attention were present at both primary and secondary level, and whether lower attaining pupils experienced more individual attention.

1.7.2.3. Pupil active involvement with teacher. Larger classes can lead to pupils having a passive role in class. Research in the UK found that children in large primary classes were more likely to engage in passive behavior, listening to the teacher, while in smaller classes pupils were more likely to interact in an active, sustained way with teachers (Blatchford et al., 2005). This can be considered another aspect of classroom engagement. This was a consistent finding at both the beginning and end of the primary stage (4/5 and 10/11 years). In Australia, Bourke (1986) found more student questions to teachers in large classes but these were mostly requests for clarification or for other help from the teacher. In this study we wanted to check whether there was more active involvement with the teacher through primary and secondary education, and whether this varied by pupil attainment level.

1.7.2.4. Easier classroom control and management – dealing with negative behavior. A number of studies have reported that pupil discipline tends to be more difficult in large classes and more of an intrusion into the teaching and learning process (e.g., Blatchford & Mortimore, 1994; Glass et al., 1982). In contrast, smaller classes tend to be quieter and more easily managed. Bourke (1986) found more non-academic procedural arrangements were necessary in large classes.

It was expected that there would be less need for critical comments from teachers in smaller classes, reflecting fewer difficulties in classroom management. In general, it might be anticipated that lower attainers would be more off task, and so may be expected to receive more critical comments from the teacher, but it was not clear whether this would be affected by size of class.

1.8. Summary of predictions

1.8.1. Pupil classroom engagement

For ease of reference in the results section, each research prediction will be numbered. It was predicted that in larger classes there would be less on task (1) and more off task behavior (2) and that this latter result would be particularly seen through distracted, passive forms of off task behavior (3). Predictions regarding attainment group and primary vs. secondary were less clear cut but generally it was anticipated that effects would be most obvious for low attaining (4) and primary aged students (5).

1.8.2. Teacher–pupil interactions

It was predicted that in larger classes there would be less teaching overall (6), less individual attention from teachers (7), a less active role in interactions with teachers (8), and more teacher talk about negative behavior (9). Predictions regarding attainment group and primary vs. secondary were not clear cut but generally it was again anticipated that effects would be most obvious for low attaining (10) and primary aged students (11).

2. Method

2.1. Research design

There may be an obvious attraction to the adage 'to understand something, change it'. The strength of the experimental design used in the STAR project is that by randomly allocating teachers and pupils to classes of a different size it is in theory easier to draw unambiguous conclusions concerning the causal role of class size. However, it is not always appreciated that there can be theoretical and practical difficulties with experimental approaches to class size effects (Goldstein et al., 1998, see also Grissmer, 1999; Hanushek, 1999; Mitchell, Beach, & Baduruk, 1989; Prais,

1996). It can more valid to seek better understanding of the effects of class size differences by measuring and examining relationships between class size and classroom processes as they occur naturally in schools, and to make adjustments for possibly confounding factors such as prior pupil attainment, gender, special educational needs and so on. A naturalistic design can be useful in addressing policy issues in that it is more 'authentic', and reflects adjustments and processes as they occur under normal circumstances. It can allow modeling of effects in the real world rather than relatively artificial comparisons. It can also be used examine effects across the full range of class sizes, not just a few selected sizes. This can be important for policy recommendations, for example, if there are certain class sizes, or class sizes below or above a certain number, which have stronger effects.

The overall strategy adopted in this study was therefore to randomly select participating schools, measure effects of naturally occurring differences in class sizes using measures carefully drawn up on the basis of previous research and pilot work, and control statistically for factors likely to interact with or confound any class size effect on classroom processes.

2.2. Sample

2.2.1. Schools

Results come from a large scale study of the deployment and impact of support staff (called the DISS project) in primary and secondary schools in England and Wales (see Blatchford et al., 2009). Systematic observations were carried out over 2005/6 in 49 schools. These schools were chosen at random from a national survey as part of the DISS project and they then agreed to field visits by researchers. The schools did not differ from the non-participating schools on main characteristics such as percentage of pupils for whom English was an additional language (EAL), percentage of pupils with Statements of Special Educational Needs (SEN), number of pupils, and whether in urban vs. rural areas. There were 27 primary schools and 22 secondary schools. Two year groups were generally observed in each school, either Year 1 and Year 3 (5/6 and 7/8 years) or Year 7 and Year 10 (11/12 and 14/15 years). Observations were conducted in 88 classes.

2.2.2. Characteristics of pupils

Observations were conducted on a sample of eight pupils in each class. Pupils were classified into three attainment groups - low, medium or high attainment groups - on the basis of teacher ratings. Attainment levels were school based and so there were slight fluctuations between schools in numbers of pupils in the three attainment groups. Alternative, continuous measures of attainment (stemming from Government assessments) were also available for just over half of the sample and analyses showed a fair degree of overlap in attainment and teacher rating measures (correlation of 0.5). As we shall see, the regression analyses of observation outcomes, to be presented in the results section, were redone with the continuous attainment measure also included, and models were virtually unaltered, indicating that teacher ratings of pupil attainment were a valid measure.

In addition to the number of pupils and prior attainment, the other variables listed in Table 1 were also included in the analyses to make sure they did not account for the effects found – see below for details on the statistical analysis. There were 686 pupils observed in total. Details of the sample are given in Table 1. Information was obtained from Government data sets or from schools themselves.

Visits lasted 4 days except when observations were only possible in one year group (such as infant or junior schools) – they then lasted 2 days - and observations were made in mathematics, English, science and Welsh lessons.

2.2.3. Class size

Class size might seem to be an obvious and easily available measure, but the number of children actually in the class at any time may to some extent be different to the number according to the class register; children may be away, for example. The view taken here is that the class size experienced by a student at the time of each observation (what we have called the 'experienced' class size) is the classroom contextual unit most likely to be connected to moment by moment classroom interactions and pupil engagement in lessons. However, in the UK at least there is in practice little difference between registered and experienced class sizes. Pupil Teacher Ratios (PTRs) are different from, and more limited than, class size, because they take no account, for example, of teacher non-contact time. At the time of each observation (i.e., each ten-second time interval) a note was therefore made of the number of children in the classroom. The distribution of class sizes was fairly even with no evidence of outliers, either large or small. For primary classes the mean was 23 with a standard deviation of 6. The median was 23 with an inter-quartile range of 20–27. For secondary classes the mean was 18 with a standard deviation of 7. The median was 19 with an inter-quartile range of 13–24.

Table 1
Systematic Observation Component: Characteristics of Pupils.

Characteristic	Category	Number	Percentage
Year	1	200	29%
	3	183	27%
	7	152	22%
	10	151	22%
Gender	Female	335	49%
	Male	351	51%
SEN status[a]	None	319	55%
	School Action	141	24%
	School Action +	57	10%
	Statement	68	12%
Attainment group	Low	123	21%
	Medium	241	41%
	High	221	40%

[a] 'School Action', i.e., requiring provision different from, and additional to, other pupils; 'School Action Plus', i.e., receiving help from sources external to the school; 'Statemented', i.e., with more severe or complex needs that require exceptional provision.

2.3. Observation system

The observations provided a moment by moment description of each pupil's behavior. The basic principle was to observe when classroom-based activities took place, and to provide a representative and systematic account of pupils' behavior. Observations were conducted on each child in turn in blocks of 10 ten-second time intervals, with gaps of twenty seconds between observations to allow recording of what took place in the previous ten seconds. There were 34,420 ten-second observations in total.

The observation categories were devised on the basis of well established systems, as used in the CSPAR and SPRinG studies (e.g., Blatchford et al., 2005; Blatchford, Baines, Rubie-Davies, Bassett, & Chowne, 2006). It provided a quantitative account of pre-specified categories of behavior, including time pupils spent in three social 'modes' - with teachers, with other children and when not interacting. Within each of these three 'modes' were categories that covered work, procedural, social and off task activity. The categories referred to the 'target' child; teachers, support staff and other children were observed only when they came into contact with them. The schedule employed a form of predominant activity sampling with those behaviors selected within sets of behaviors (e.g., social modes) occurring for the longest period within the ten-second interval. Behaviors within sets were mutually exclusive and exhaustive. Not all categories feature in this paper and brief definitions of categories used in this paper are as follows:

2.3.1. Teacher/pupil interaction

2.3.1.1. Teacher to child talk

2.3.1.1.1. Teacher 'teach'. Behavior directly concerned with the substantive content of subject knowledge, i.e. communicating concepts, facts or ideas by explaining, informing, demonstrating, questioning, suggesting.

2.3.1.1.2. Teacher 'on task'. As Teacher '*teach*' plus *task preparation*, i.e., contacts concerning the organization and preparation of children's task activities.

2.3.1.1.3. Dealing with negative behavior. When the teacher had to correct the target child or a group within which the target child belonged. The category would not have included simple academic disagreements over an answer from a pupil, but rather times when the teacher deliberately dealt with a child was considered to be off task, behaving inappropriately or misbehaving.

2.3.1.2. Child role

2.3.1.2.1. Child 'focus'. Target child is the focus of the teacher's attention, and this could be in the context of one-to-one, group or whole class sessions.

2.3.1.3. Child to teacher interaction

2.3.1.3.1. Child active interaction with teacher. The sum of *initiate* (begins an interaction), *respond* (responds to an adult initiation), and *sustains* (i.e., interaction extends over a ten-second time interval).

2.3.1.3.2. Child on task to teacher. All child behaviors in contact with adult that are concerned with work, including listening to the teacher teach.

2.3.1.3.3. Child off task to teacher. Child behavior when in contact with the adult obviously inappropriate or unrelated to situation either passively (e.g. not attending) or actively (e.g., talking).

2.3.1.4. Pupil–pupil interaction

2.3.1.4.1. Target and child on task. All contacts with other children that are concerned with work and allocated tasks.

2.3.1.4.2. Target to child off task. Behavior with other children that is deliberately off-task; it would include 'mucking about' and fooling around and times the target child is aggressive (verbally or physically) towards other child(ren).

2.3.1.5. Individual behavior/not interacting

2.3.1.5.1. Individual on task. Target child is involved in own work activity.

2.3.1.5.2. Individual off task (active). Target child focuses on something other than task in hand.

2.3.1.5.3. Individual off task (passive). Target child is disengaged during task activity, for example, daydreaming.

2.3.1.6. Computed categories

2.3.1.6.1. Child on task. Total on task behaviors, i.e., behaviors related to the substantive nature of allocated work or preparation for the work across the three social modes, i.e. child to teacher on task, target and child on task, and individual on task.

2.3.1.6.2. Child off task. Total off task behaviors, i.e., all off-task behaviors in the three social modes, i.e., child to teacher off task (active or passive), target to child off task, and individual off task (active and passive).

2.3.2. Observers

There were two observers. They were experienced researchers who were familiar with working in schools, able to explain the research and put teachers and pupils at their ease. The basic aim was to avoid passing judgments, and to use the schedule as intended. One observer had extensive experience of field work in schools involving systematic observations. The other observer had initial training in which they were provided with an observation manual of categories, conventions and procedures, as well as tips acquired during previous use. Conventions were discussed and there was work on videotapes, accompanied by periodic checks of accuracy and understanding of how to use categories. This was followed by four day's observation of the same pupils in a class, and follow up sessions to iron out any difficulties.

2.3.3. Reliability checks

Reliability was addressed by calculating agreement between two independent observers for a sub-set of observations. Kappa coefficients (which correct for chance agreement) for the main

sets of mutually exclusive categories, e.g.,teacher-child talk, 'child role', child to teacher interaction, not interacting', were all high, with reliability coefficients greater than 0.80. Kappa for pupil–pupil interaction was 0.77.

2.4. Statistical methods and analysis of systematic observation data

The limitations of much previous research, that makes use of observational data, are not always appreciated. A feature of the analysis of the observation data in this paper was the way that it was conducted with the 10-second observation interval as the unit of analysis. This allows a greater accuracy and flexibility than simple, but more commonly used, total frequencies of behaviors for each pupil. In particular it provides the basis for powerful analyses of the co-occurrence of behaviors. This kind of analysis is not possible when simple totals for each pupil are used. A further feature of this observation study, in contrast to previous research, is that it used multilevel statistical models. These were required as it is likely that observations from pupils in the same class will be more similar than two observations from pupils in different classes. Similarly, two observations from the same pupil are more likely to be similar than two observations from differing pupils. Therefore, the observations cannot be regarded as independent of each other, and so multilevel statistical methods (Goldstein, 1995) are required. Failure to allow for the structure of the data can lead to misleading parameter estimates, and also an exaggeration of the significance of the results (Paterson & Goldstein, 1991).

Three level models were used for the analysis, with individual observations contained within pupils, contained within classes. A potential fourth level (the observation sheet) was also considered but the results had more stability when there were only three levels, and so this option was not used. School level was not included in the analysis because it did not add much over the class level. There were schools with only one class and exploratory analysis with school included suggested models did not support both class and school levels. Potentially important background characteristics like pupil attainment levels and eligibility for free school meals, which might be connected to school differences, were in any case measured at the pupil level.

As all observation outcome variables were binary in nature, in the sense of each either being performed or not being performed during one time interval, logistic regression models were used to examine the effects of the various explanatory variables on the outcomes. The following regression equation describes the analysis used for the modeling, simplified to include only one predictor variable.

Let y_{ijk} be the observed binary response (0, 1) of observation i from pupil j in class k, and let the probability of a particular outcome being observed be denoted by π_{ijk}, where $\pi_{ijk} = [\Pr(y_{ijk}) = 1]$.

The basic form of the regression equation used in the analysis for a single explanatory variable x is given by:

$$\mathrm{logit}(\pi_{ijk}) = ß_0 + ß_1 \cdot X_{ijk} + u_k^{(3)} + u_{jk}^{(2)}$$

Where $u_k^{(3)}$ is the random effect at the class level, and $u_{jk}^{(2)}$ the random effect at the pupil level. All random effects are assumed to be normally distributed as follows:

$$u_k^{(3)} \sim N\left(0, \Omega_k^{(3)}\right), u_{jk}^{(2)} \sim N\left(0, \Omega_{jk}^{(2)}\right),$$

The regression models were fitted using the MLwiN software package (Goldstein, Rasbash, Plewis, Draper, Browne, Yang, Woodhouse & Healy, 1998).

The main explanatory variable of interest was class size, which was centered before analysis. The effect of class size and other explanatory factors thought likely to influence the observation outcomes were examined jointly. The aim was to estimate the effect of class size, controlling for the effects of the other explanatory factors. In addition to class size, the following variables were also included in the analyses:

- Pupil level of attainment at the start of year (low, medium or high)
- Special Educational Needs (SEN) status of pupils (none, school action, school action plus/SEN)
- Gender
- Number of teachers
- School Subject
- Support staff presence

The advantage of including these variables in the analysis is that the effect of class size on the outcomes is adjusted to account for any effects that these variables might have on the outcomes. This therefore tells us whether size of class has an independent effect. Pupil attainment level, SEN status and gender were treated at the pupil level, whilst class size, number of teachers, school subject and support staff presence varied at the individual observation level.

In the interests of space we do not present the full regression results, or variance components. There are a large number of outcomes included in this paper and presenting the full models would greatly increase the size of the tables. The focus was only on fixed effects (specifically class size). Although we allowed for the multilevel structure in the data, the random effects were of less interest in this instance. In addition, given the binary nature of the outcome, the Level 1 variation is fixed, and so the random effects cannot be so easily interpreted as might be the case with a continuous outcome. All covariates were included in all models and were treated as fixed factors; no additional error terms were therefore present in the models. No additional class level random effects were added for any of the covariates, including class size, as these were not found to significantly improve the fit of the model.

The results for other covariates are not reported. It is not implied that class size is the only, or even the most important, influence. Pupils with Special Educational Needs, for example, were less likely to be 'on task'. But this result is not surprising and, as class size is the main focus of the paper, this and other similar results are not reported in the paper.

The interaction between class size and pupil attainment was assessed in every model. This examined if the effect of class size varied for pupils of different attainment levels. Where significant interactions were found, the results are reported separately for each subgroup; otherwise the results are presented for all pupils combined. There will be no additional error term required by including a cross-level interaction between size and attainment.

Quadratic and cubic terms for class size were included in the model when suggested by the data. If the inclusion of a cubic and quadratic term did not significantly improve the fit of the model, then these were not included, so as not to overly complicate the model.

For each of the outcomes, the analyses were performed separately for primary and secondary schools.

For the analyses reported in this paper outcomes are considered in relation to all observations in the analysis, rather than just those from a sub-set. The result of this is the analyses effectively examine the effect of class size on outcomes as a proportion of all observations.

In the tables below (in the results section), the size of effect is in terms of odds ratios, and these indicate the change in the odds of the outcome occurring for a 5 pupil increase in class size. When a simple linear relationship is used, an odds ratio greater than one means that class size was positively associated with the outcome occurring, and an odds ratio less than one means class size was negatively related to the observation outcome. For example, an odds ratio of 1.25 means that a five pupil increase in class size was associated with the odds of an observation outcome increasing by 25%. Also given are 95% confidence intervals for each odds ratio, as well as *p*-values indicating the significance of each result. Note that when a non-linear relationship was found the odds ratios do not have such an easy interpretation, and the shape of the relationship is best illustrated by the graphical plots.

Odds ratios are presented in preference to regression coefficients as they are more interpretable. In principle, beta values can be obtained for logistic regression but they are slightly less reliable than for linear regression. Moreover, a beta value would assume a linear effect of class size, which would not mirror a lot of the findings.

It is also possible in the graphs to read off the probability of a behavior occurring for any given size of class, e.g., to compare the probability of a behavior occurring in a large class of 30 vs. a relatively small class of 15 (Note that these class sizes are selected because they represent large and small classes not because they are seen to be particularly educationally significant). These probabilities are useful, and easily interpretable, i.e., it can be taken as the occurrence of any given behavior occurring, as a proportion of the total number of observations. For example, a probability of 0.8 for an observation outcome occurring at a class size of 30 means that the outcome occurred in 80% of all observations. However, some caution should be exercised when interpreting the probabilities, as these will be dependent on the other terms in the models (e.g. pupil level of support, SEN status etc.). The graphs show the probabilities for a female pupil of medium attainment, with no SEN and no support, in English lessons.

2.5. Checks on possible confounding factors

As has been explained, in this study a non-experimental, correlational design was used within which potentially confounding factors were factored into the regression analyses. For any potentially confounding factor to be important it would need to be related to class size and the observation outcomes. Factors included were: prior attainment, level of SEN, gender, number of teachers, and school support staff presence. If relationships are still found between class size and the observation measures, this suggests that class size has an independent effect.

In addition, analyses were conducted to clarify whether there was a relationship between class size and pupil attainment, which might confound relationships between class size and observation outcomes. These showed that children of different attainment levels (as rated by the teachers) in Years 1 and 3 did not tend to be in smaller classes while lower attaining pupils at secondary level (i.e., in Years 7 and 10) did tend to be in smaller classes. However, attainment level was included in the regression analyses (see below), and so relationships between class size and outcomes would have been adjusted for this as a possible explanatory variable. As a further check, the regression analyses of observation outcomes, to be presented in the results section, were redone with the continuous attainment measure also included, and models were virtually unaltered, further indicating that pupil attainment was not an extra factor accounting for relationships found.

3. Results

We first present the frequency of each of the behaviors in Table 2, expressed in terms of the numbers of observations in each category, plus the percentage of the total observations that this represents.

It can be seen that pupils at both primary and secondary stages spent a large portion of their time engaged in on task activities. The bulk of this involved on task pupil to teacher behavior, which for the most part involved listening to the teacher. In contrast, off task behavior occupied less time, though more at secondary than primary (17% vs. 11 % of all observations). In the case of the teacher to pupil categories, in just under half of observations pupils were involved in teaching interactions – in which teachers covered the substantive content of school subjects. By contrast, the three categories which denoted individual interactions with target pupils were infrequent, only occurring in between 2 and 6% of all observations. In particular, teachers rarely criticized individual pupils, as a proportion of all observations.

3.1. Classroom engagement

3.1.1. Total pupil on task

The first outcome was total on task behavior (i.e., behaviors related to the substantive nature of allocated work or

Table 2
Frequencies and percentages of occurrence of each behavior.

Behaviour Group	Behavior	Primary Number (%)	Secondary Number (%)
On and Off task Behavior[a]	Total On task	15,269 (86%)	13,262 (80%)
	On task - pupil to teacher	10,286 (58%)	7983 (48%)
	On task – pupil–pupil	1835 (10%)	1846 (11%)
	On task – not interacting	3501 (20%)	3751 (23%)
	Total Off task	1931 (11%)	2765 (17%)
	Off task – pupil to teacher	997 (6%)	1078 (6%)
	Off task – pupil–pupil	679 (4%)	1345 (8%)
	Off task – not interacting active	373 (2%)	573 (3%)
	Off task – not interacting passive	749 (4%)	668 (4%)
Teacher–pupil interaction	Teacher teach	7524 (42%)	7288 (44%)
	Pupil focus of Teacher	840 (5%)	942 (6%)
	Pupil active interaction with teacher	791 (4%)	910 (6%)
	Teacher deal with negative behavior	276 (2%)	423 (3%)

Note. All observation measures are considered in relation to all observations in the analysis, and not just those from a sub-set. The analyses therefore examine the effect of class size on outcomes as a proportion of all observations.

[a] The on and off task sub-categories slightly exceed the total, as in some observations more than one sub-category was coded. Total on and off task behaviors were calculated by summing time intervals in which these behaviors occurred, whether or not they occurred more than once.

preparation for the work across the three social modes - child to teacher on task, target and child on task, and individual on task). The results addressed research predictions 1–5.

The results (See Table 3) indicated that for primary schools there was no significant interaction between attainment group and class size, and that the effect of class size did not therefore vary by attainment group. There was a statistically significant association between number of pupils and on task behavior ($p < 0.05$). A higher number of pupils were associated with a decreased occurrence of on task behavior.

The results for secondary pupils indicated that the effect of class size on on task behavior varied by attainment group. There was no significant effect of class size for pupils in the medium and high attainment groups. However, for pupils in the low attainment group, a larger number of pupils was associated with a decreased occurrence of on task behavior. A five pupil increase in class size was associated with the odds of on task behavior decreasing by almost a quarter. Looking at Fig. 1 shows that the difference between 30 and 15 is about 78% vs. 88%, i.e., a 10% difference for low attaining pupils – a larger difference in comparison to primary schools. This shows that there was an interaction effect for research predictions 4 and 5; the effect was most obvious for low attaining pupils at secondary school only.

3.1.2. Total pupil off task

The relationship between the number of pupils and occurrence of total off task behavior are also shown in Table 3. In general, as expected, the results were the converse of those for on task behavior. (They were not exactly related because the two categories added together were not exhaustive of all behaviors; i.e., pupils could be engaged in behaviors other than on and off task behavior.)

The results indicated that the effect of class size on off task behavior varied for pupils of differing attainment. For primary schools there was an increase in off task for low and medium attaining pupils. For the low attainment group, a five pupil increase in class size was associated with the odds of off task increasing by 11%. There was no significant effect of class size for the high attainers. The results for primary schools are illustrated in Fig. 2.

In line with results for on task behavior, the results for secondary schools indicated a highly significant effect of class size for low attaining pupils only. A five pupil increase in class size was associated with the odds of off task behavior increasing by 40% for this group. Looking at this in terms of probabilities of occurrence with 15 vs. 30 in a class (see Fig. 3) shows that about 0.26 of observations were off task for a class size of 30, but only 0.11 of observations were off task with 15 in a class. This is the difference between 26% and 11% of all observations. Low attainers therefore spend more than twice as much time off task in large vs. small classes, a sizeable difference.

There was no strong evidence of an effect of class size for either the medium or high groups, although there was slight evidence that off task was less likely in larger classes for the high attainers. However, this result was not quite statistically significant ($p = 0.07$).

Results for on and off task behavior were therefore in one sense the opposite of that predicted in research prediction 5, i.e., the effect was stronger for secondary aged pupils.

3.1.3. Separate categories of on and off task behavior

Separate analyses were conducted on each of the seven individual on and off task categories, i.e. pupil to teacher on task, pupil–pupil on task, and not interacting on task, and pupil to teacher off task, pupil–pupil off task, and not interacting off task active and passive. In the interests of space these results are not reported in full here. There did not appear to be a stronger relationship between class size and passive, compared to active, forms of off task behavior (when not interacting), contrary to research prediction 3. Pupil to teacher on and off task behaviors were the most consistent with the total on and off task behavior results. This is not surprising as pupil to teacher behavior was the most common of the three

Table 3
Class size and observation measures.

	School type	Attainment group	Term	Odds Ratio (95% CI)	*P*-value
Class size and total on task behavior	Primary	All pupils	Linear	0.88 (0.82, 0.94)	0.04
	Secondary	Low	Linear	0.78 (0.71, 0.86)	<0.001
		Medium	Linear	1.01 (0.89, 1.15)	0.86
		High	Linear	1.12 (0.92, 1.38)	0.26
Class Size and Total Off Task Behavior	Primary	Low	Linear	1.11 (1.02, 1.21)	0.01
		Medium	Linear	1.25 (1.07, 1.47)	0.005
		High	Linear	0.91 (0.77, 1.06)	0.23
	Secondary	Low	Linear	1.41 (1.27, 1.57)	<0.001
		Medium	Linear	0.96 (0.84, 1.10)	0.57
		High	Linear	0.81 (0.64, 1.01)	0.07
Class Size and Teacher 'Teach'	Primary	All pupils	Linear	1.05 (0.98, 1.12)	<0.001
			Quadratic	0.95 (0.92, 0.98)	
	Secondary	All pupils	Linear	1.08 (1.02, 1.15)	0.01
Class Size and Pupil Focus of Teacher's Attention	Primary	All pupils	Linear	0.73 (0.68, 0.79)	<0.001
	Secondary	All pupils	Linear	0.72 (0.66, 0.80)	<0.001
Class Size and Active Interactions with the Teacher	Primary	All pupils	Linear	0.73 (0.70, 0.87)	<0.001
	Secondary	All pupils	Linear	0.79 (0.71, 0.87)	<0.001
			Quadratic	1.05 (0.99, 1.11)	
Class Size and Teacher Dealing with Negative Behavior	Primary	Low	Linear	1.31 (1.01, 1.70)	0.04
		Medium	Linear	1.53 (1.13, 2.07)	0.01
		High	Linear	0.82 (0.58, 1.14)	0.24
	Secondary	Low	Linear	1.16 (0.99, 1.37)	0.07
		Medium	Linear	0.80 (0.65, 1.00)	0.04
		High	Linear	0.78 (0.54, 1.13)	0.19

'social modes' (i.e., more than when with peers or when not interacting). Interestingly there was a tendency for both pupil to pupil on and off task behaviors to increase with size of class, indicating that pupils tended to interact more with each other, in both on and off task ways, as the number of pupils in the class increased.

3.2. Teacher pupil interactions

3.2.1. Teacher teach

Similar analyses were performed to examine the effect of class size on the total amount of teacher 'teach', ie., teacher interactions directly concerned with the substantive content of subject knowledge (see Table 3).

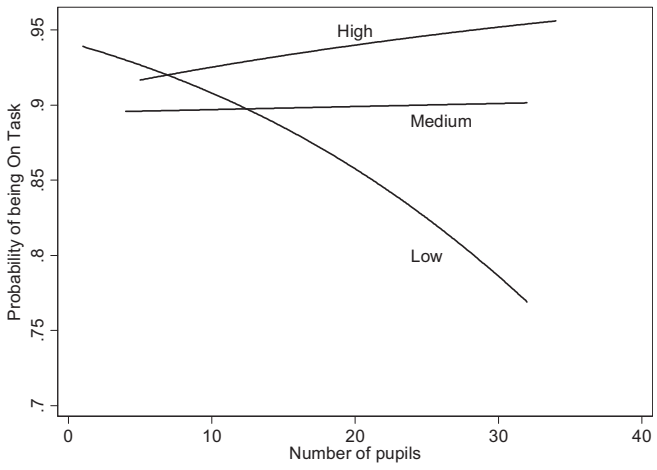

Fig. 1. Class size and total on task behavior (Secondary).

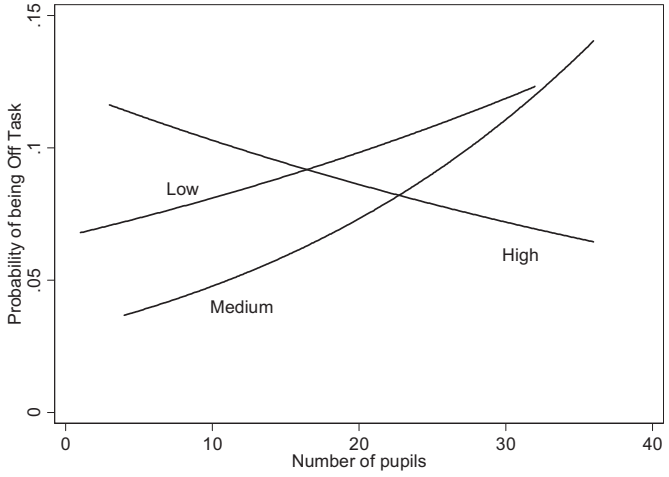

Fig. 2. Class size and total off task behavior (Primary).

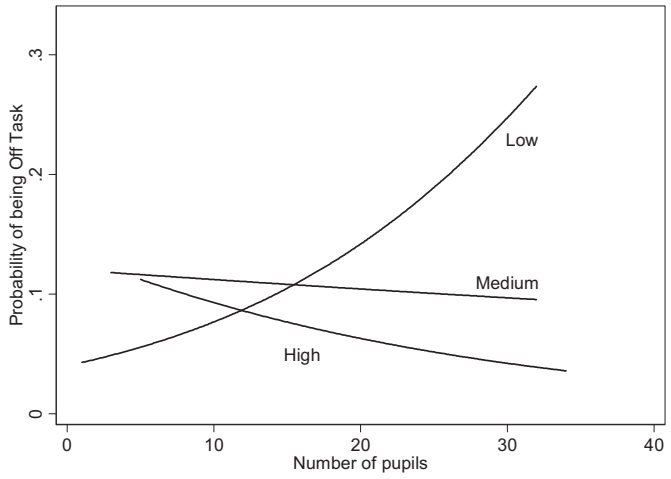

Fig. 3. Class size and total off task behavior (Secondary).

There was no evidence of an interaction between class size and attainment group for either school phase. There was a significant effect of class size on teacher teach in primary schools; there was generally more teacher teach in larger classes, although this tailed off for the largest classes.

There was also a significant effect of class size for secondary schools (see Table 3), and again there was a positive association between class size and the amount of teacher teach (see Fig. 4). An increase in class size of 5 pupils was associated with the odds of teacher teach increasing by 8%. Roughly speaking the difference between 30 and 15 in the class means a difference between 52% and 45% of all observations – i.e., in the smaller class there is around 7% less 'teach' occurring. These results were the opposite of that expressed in research prediction 6.

3.2.2. Pupil focus of teacher

There was a highly significant association between class size and pupil being the focus of a teacher's attention (see Table 3). Though 'focus' did not occur very frequently, it noticeably increased as class size decreased. The results were displayed in graphical form in Fig. 5. The difference between 30 and 15 in the class represented a difference of about 7% vs. 3% of all observations, i.e., focus was more than halved in a large vs. a small class.

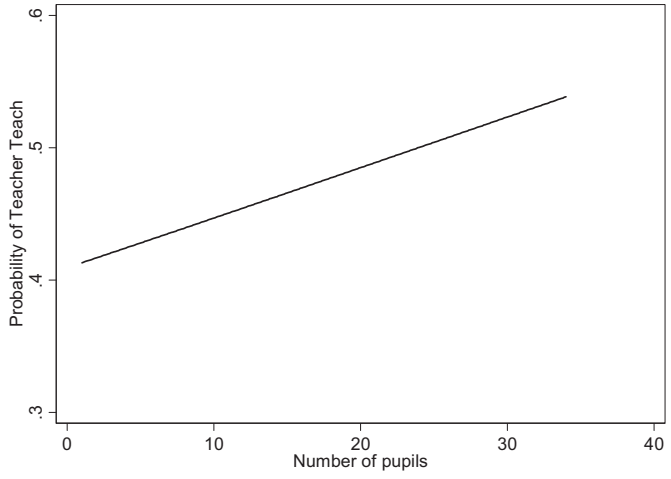

Fig. 4. Class size and teacher 'teach' (Secondary).

There was a similar, statistically significant effect of class size at secondary. An increase in class size of five pupils resulted in the odds of pupil being the focus of a teacher's attention decreasing by about a quarter.

There were no interactions with pupil attainment at either primary or secondary.

These results for pupil focus were in line with research prediction 7, though there were no differences between attainment groups or primary vs. secondary (and therefore this is not in line with research predictions 10 and 11).

3.2.3. Pupil active interaction with teacher

The effects of class size and the amount of pupil active interaction with the teacher are shown in Table 3. As with the other results, the figures show the change in the odds of the outcome occurring for a 5 pupil increase in class size.

At primary level, there was a significant negative effect of class size, and this result is illustrated in Fig. 6. There was no interaction with the attainment level of pupils; the effect was similar for all three groups. The difference between 30 and 15 in the class was about the difference between 2% and 6% of all observations. Though not frequent, there is about three times more active interaction in small classes.

For secondary schools there was also a highly significant effect of class size on the occurrence of active interaction with the teacher. As in primary schools there was less active interaction with the teacher in larger classes. These results for pupil active involvement with the teacher were in line with research prediction 8, though there were no differences between attainment groups or primary vs. secondary (i.e., not in line with research predictions 10 and 11).

3.2.4. Teacher dealing with negative behavior

The effect of class size on teachers dealing with negative behavior is shown in Table 3.

For both school phases the effect of class size varied for pupils with different attainment.

For primary schools there was significantly more teacher dealing with negative behavior in larger classes for low and medium attaining pupils, but no significant effect for high attainers. For low attaining pupils the odds of a teacher dealing with negative behavior increase by about 30% for an increase in class size of five pupils. An illustration of the primary results is given in Fig. 7.

The results for secondary schools indicated that there was significantly less of the teacher dealing with negative behavior in larger class for medium attaining pupils. Conversely, there was slight evidence that there was more dealing with negative behavior in large classes for low attaining pupils, although this result was not quite statistically significant. There was no significant effect for high attainers. These results for dealing with negative behavior were in line with research prediction 9, but as with on and off task behavior, showed an interaction between pupil attainment

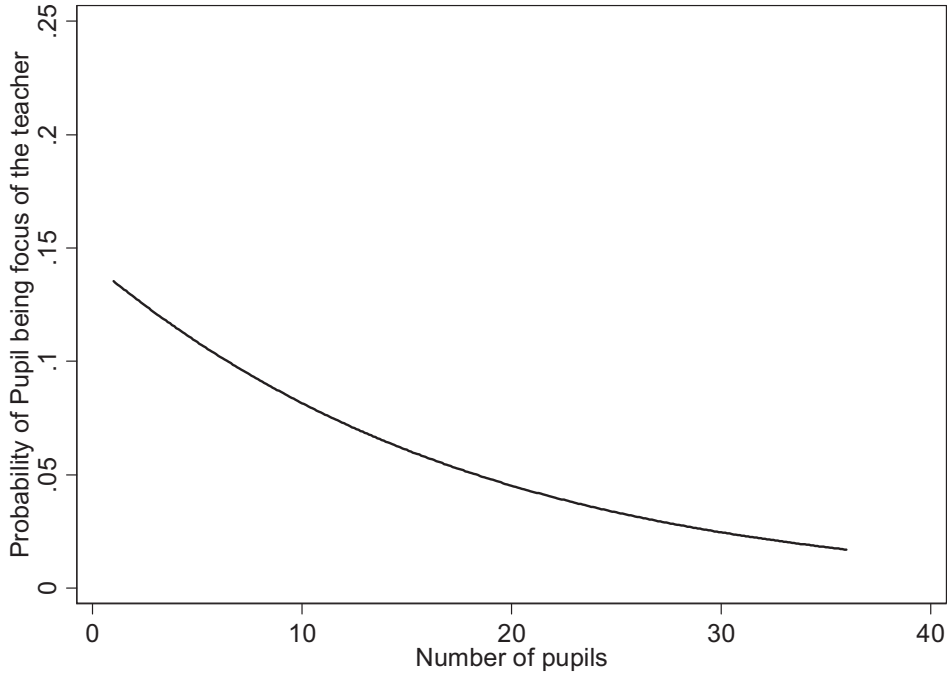

Fig. 5. Class size and pupil focus of teacher's attention (Primary).

level and primary vs. secondary, i.e., effects were most marked for low attaining pupils at secondary level (i.e., in line with research prediction 10 but not 11).

For ease of reference a summary of results is given in Table 4.

4. Discussion

This study aimed to extend previous research by comparing effects of class size on pupil classroom engagement and teacher–pupil interaction, and by examining if effects varied by pupil attainment level and between primary and secondary school years. The study was based on a naturalistic, non-experimental design and so associations cannot be taken strictly as evidence of causal direction. Nevertheless the analysis controlled for possibly confounding factors, such as pupil attainment level, SEN status, gender, the number of teachers, school subject and the presence of support staff, and there was no evidence that these accounted for effects found. There was no evidence that the participating schools differed from non-participating schools. This further suggests relationships between class size and observation outcomes reported in this paper are not explained by any other factor.

4.1. Teacher–pupil interactions

Perhaps the clearest result to emerge from this study is the way that class size was related to the amount of individual contact with pupils. This was evidenced through two particular types of behavior: 1. times when the pupil was the focus of

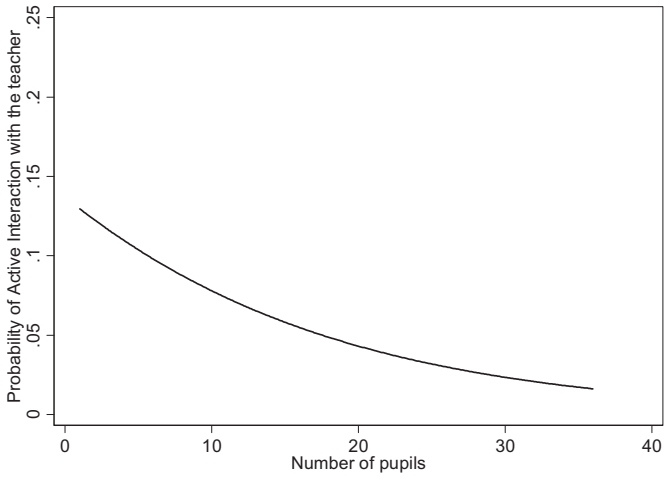

Fig. 6. Class size and active interactions with the teacher (Primary).

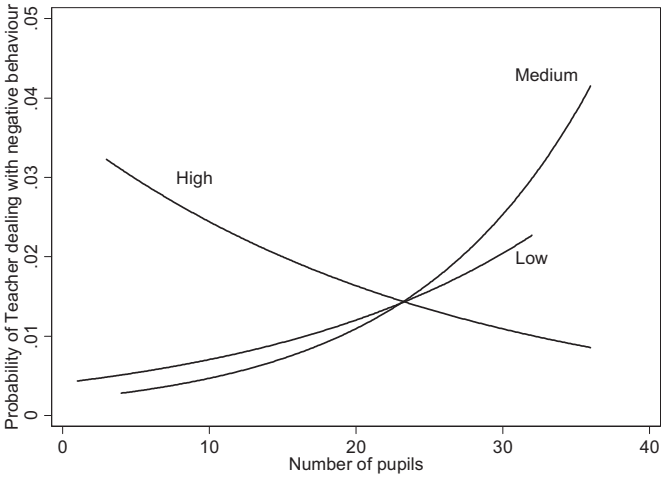

Fig. 7. Class size and teacher dealing with negative behavior (Primary).

Table 4
Summary of results on relationships between class size and observation outcomes.

Behaviour	Primary	Secondary
On and Off task Behaviour		
Total On task	All groups LESS	Low LESS
Total Off task	Low and Middle MORE	Low MORE
Teacher–pupil interaction		
Teacher teach	All groups MORE	All groups MORE
Pupil focus of Teacher	All groups LESS	All groups LESS
Pupil active interaction with teacher	All groups LESS	All groups LESS
Teacher deal with negative behaviour	Low/Middle MORE	Low MORE ($p = 0.07$); middle LESS

a teacher's attention and 2. times when they were engaged in active interactions with their teachers, i.e., beginning, responding to, and sustaining interactions with them. The converse also applies – as class sizes became smaller there were more times when pupils were the focus of a teacher's attention, and more times when they were engaged in active interaction with teachers. This effect was found for all groups at both primary and secondary levels and there was little evidence that the relationship between class size and teacher pupil interaction varied by attainment groups. Though these behaviors were not frequent, when seen as a percentage of all observations, there was between two and three times more of these behaviors in smaller classes of 15 compared to larger classes of 30. This appears to support predictions and previous research (see Introduction) but contrary to expectations shows that effects of class size extend beyond primary school into secondary school.

4.2. Amount of teaching

Another result to emerge from the analysis of teacher pupil interaction was the finding that the amount of teaching, i.e., talk dealing with the substantive nature of a task, through explaining or questioning etc, increased as the size of class increased. This was contrary to the predicted direction of effects and may therefore seem contradictory. It seems that pupils get less individual attention in larger classes but they also receive more of a teacher's input overall relating to educational matter. On the face of it this might seem to mean that larger classes advantage pupils. However, the finding probably means that pupils as a whole are actually receiving more of a teacher's delivery of a lesson in whole class contexts. This is confirmed by other results from the same study, not reported here, which showed that for primary and secondary schools together there was more whole class teaching in larger classes. Putting these two main results together therefore suggests that in smaller classes pupils get more individual attention, while in larger classes they spend more time listening to the teacher talk to the whole class. They are perhaps getting more educational input, but this is at the expense of it being largely passive and part of a large group.

4.3. Classroom engagement

The other main set of results reported in this paper concerned classroom engagement. Here we found that there was a tendency for there to be more pupil on task and less off task behavior as class sizes decreased, and conversely less on task and more off task behavior as class sizes increased. Unlike results for teacher pupil interaction, however, this was affected by the pupil's attainment group and also primary vs. secondary. While there was more on task in smaller classes in primary schools for all attainment groups, at secondary level it was only the low attainers who showed more on task behavior. For illustrative purposes we compared a large class of 30 with a small class of 15 and this showed a difference of about 10% in on task behavior for low attaining pupils. In the case of off task behavior, at primary level it was the middle and low pupils who showed most off task behavior in larger classes, but at secondary level it was again the low attainers who tended to be most affected. We found that for the low attainers at secondary level there was more than twice as much off task behavior in large classes of 30 compared to smaller classes of 15. These findings on classroom engagement and class size confirm predictions but are new in that they extend previous research into secondary and are clear that it is the low attaining pupils whose attention is most affected by larger classes.

These findings are probably connected to results on the effect of class size on teachers dealing with negative behavior. This category was coded when teachers had to correct the target child or a group within which the target child was situated, when the teacher perceived them to be off task and misbehaving. In line with results on off task behavior, we found the amount of dealing with negative behavior was affected by pupil attainment group. The clearest trend is for the low attaining group to receive more of this kind of corrective behavior, at both primary and secondary level. It therefore looks as if teachers in larger classes are responding to the increased off task behavior of low attaining pupils by seeking to control their behavior.

These results on on and off task behavior are significant for educational effectiveness because they show that the problem of large classes, especially in older pupils, is particularly marked for the pupils who are already attaining at lower levels, and that it, in turn, is associated with teachers seeking to control this kind of behavior. It is easy to see how these two kinds of behaviors can become self reinforcing, exacerbating each other and making the situation worse. In contrast, smaller classes seem to allow an environment in which low attainers are not only less off task but are less likely to receive corrective talk from their teachers. This appears to be a more productive educational environment.

4.4. Different types of on and off task behavior

Separate analyses of different forms of on and off task behavior indicated that pupil to teacher on and off task behavior was most consistent with the results for total on and off task behavior. For the most part this involved listening to the teacher talk (active forms of interaction were relatively uncommon). So

overall it looks as if in larger classes low attaining pupils in particular are less likely to be on task and this is predominantly because they are less likely to be attending (or responding appropriately) to the teacher. It was interesting to see that pupil to pupil on and off task behavior tended to increase with size of class, at both primary and secondary level. It seems likely that with more pupils in the classroom there is less of the teacher's individual attention available and this presumably means that pupils come to interact more with each other instead, in both acceptable and unacceptable ways.

4.5. Primary vs. secondary

As was said in the introduction, relatively little is known about the effects of class size on classroom processes in older school pupils. A main aim of the study was therefore to see whether effects of class size continued into the secondary stage. In line with results on class size effects on attainment outcomes, it was predicted that effects on the two main sets of behavior - teacher pupil interaction and classroom engagement – would be most prevalent at primary level. But against expectation we found that effects actually extended into secondary schools and did not weaken, though in the case of classroom engagement effects were most marked for low attaining pupils. This therefore extends previous research findings which have reported only on primary aged pupils. The effect on individualized attention in particular, appears to be a robust finding that extends right through the school years.

4.6. Thresholds

As we saw in the introduction, it is often assumed that class sizes need to fall below a certain number (the number of 20 in a class is often cited) before they can have an impact on educational outcomes. In the present study an alternative approach was used and we were able to examine the effects of class size across the full range of class sizes, rather than presuppose class sizes likely to be important. Although class sizes of 30 and 15 were compared for illustrative purposes, there was not a clear and consistent picture regarding class sizes below or above which effects were most evident. Whilst it is recognized that the threshold debate has mostly addressed academic outcomes, the present findings suggest that it is probably over simplistic to talk about thresholds below and above which effects across all outcomes knock in, or identify optimal class sizes in an exact way. The situation is also likely to be affected by pedagogical beliefs and practices, teachers' views about preferred class sizes, their experiences of class sizes, and what they perceive as realistically achievable. Judgments are also likely to be affected by culturally bound views about teaching and about learning (see Blatchford and Lai, 2010).

4.7. Conclusions

In this paper our purpose has been to take a focused perspective on several relatively common and low inference behaviors, amenable to systematic observation techniques, and to address these in terms of the impact of pupil attainment level and age. This is not meant to imply that the present study has addressed all the classroom processes involved. The nature of the observation method used means that the emphasis has been on the quantity of behaviors, and not qualitative differences. Moreover, it is highly likely, as argued previously (Blatchford, 2003a), that class size effects are multiple rather than singular. There are likely to be other kinds of classroom processes which are affected by class size, though many of these dimensions are likely to be far harder to measure. Anderson (2000) identified a number of these factors, for example, greater knowledge of students and more 'in depth' treatment of content in smaller classes.

Perhaps the main implication of this study is that smaller classes can benefit all pupils in terms of individual, active attention from teachers, but that the lower attaining pupils in particular can benefit from small classes at secondary level. This suggests that small classes can be a valuable educational initiative right through school, but lower attaining pupils at secondary level could particularly benefit from small classes. If placed in large classes, the evidence is that they will be more prone to go off task and teachers will have to use up more time bringing them back on task.

Small classes can therefore allow teachers to engage in more individualized teaching, and can be used as part of more differentiation of the curriculum. It is well known from research (Evertson & Randolph, 1989; Graue, Rauscher, & Sherfinski, 2008), that teachers do not always adapt their teaching to take advantage of small classes. In the CSPAR case studies, it was found that some teachers in small classes still relied a good deal on whole class teaching with very brief interactions with individuals, and did not take advantage of the possibilities of increased individualization. It is also recognized, as Galton (1998) has pointed out, that we do not have a lot of knowledge about effects of class size on teaching on which to base practical advice and build coherent pedagogies. If we are serious about implementing a more individualized pedagogy then we need to think through ways in which we can maximize opportunities for individual attention, in the context of changing classroom contexts.

Another implication of this study is the need to be aware of how pupils in large classes can drift off task through too much teacher to whole class talk, and how it is the low attainers who seem most affected. In the UK at least, there is at present in schools a large diet of passive listening to the teacher talk – a diet made worse by large classes. This suggests the value of more varied pedagogical approaches. We need to be careful not to overlook the benefits that can stem from other contexts for learning, for example, pupils learning together with a deliberate attempt to minimize the teacher's input. Indeed, there is no guarantee that smaller classes will automatically lead to more productive work in groups. Research has found that if anything there is less collaborative group work in smaller classes (Blatchford, Baines, Kutnick, & Martin, 2001). This finding was also replicated in other results from the current study, not reported in this paper, in which we found a tendency for there to be less group work in smaller classes. A lot of research from many countries has shown the case for

collaborative group work, though this needs careful development, and training for both teachers and pupils (Blatchford et al., 2006). It therefore seems that teachers should be encouraged to take up opportunities for varied pedagogical approaches in smaller classes, including collaborative learning, rather than simply increase the amount of individualized attention.

References

Achilles, C. A. (1999). *Let's put kids first, finally: Getting class size right*. Thousand Oaks, California: Corwin Press.

Anderson, L. W. (2000). Why should reduced class size lead to increased student achievement? In M. C. Wang, & J. D. Finn (Eds.), *How small classes help teachers do their best* (pp. 3–24) Philadelphia: Temple University Center for Research in Human Development.

Bennett, N. (1996). Class size in primary schools: Perceptions of head teachers, chairs of governors, teachers and parents. *British Educational Research Journal, 22*(1), 33–55.

Betts, J. R., & Shkolnik, J. L. (1999). The behavioral effects of variations in class size: the case of math teachers. *Educational Evaluation and Policy Analysis, 21*(2), 193–213.

Biddle, B. J., & Berliner, D. C. (2002). *What research says about small classes and their effects*. Part of series. In *In Pursuit of better schools: What research says*, Downloadable from http://www.WestEd.org/policyperspectivesor http://edpolicyreports.org.

Blatchford, P. (2003a). *The class size debate: Is small better?* Maidenhead, U.K. and Philadelphia, U.S.A: Open University Press.

Blatchford, P. (2003b). A systematic observational study of teachers' and pupils' behaviour in large and small classes. *Learning and Instruction, 13*(6), 569–595.

Blatchford, P., Goldstein, H., & Mortimore, P. (1998). Research on class size effects: a critique of methods and a way forward. *International Journal of Educational Research, 22*, 691–710.

Blatchford, P., Baines, E., Kutnick, P., & Martin, C. (2001). Classroom contexts: connections between class size and within class grouping. *British Journal of Educational Psychology, 71*(2), 283–302.

Blatchford, P., Baines, E., Rubie-Davies, C., Bassett, P., & Chowne, A. (2006). The effect of a new approach to group-work on pupil-pupil and teacher-pupil interaction. *Journal of Educational Psychology, 98*, 750–765.

Blatchford, P., Bassett, P., & Brown, P. (2005). Teachers' and pupils' behaviour in large and small classes: a systematic observation study of pupils aged 10/11 years. *Journal of Educational Psychology, 97*(3), 454–467.

Blatchford, P., Bassett, P., Brown, P., Martin, C., Russell, A., & Webster, R. (2009). *The deployment and impact of support staff (DISS) project. Research brief 148*. Department for Children, Schools and Families (DCSF), ISBN 978 1 84775 515 5. Available from www.schoolsupportstaff.net.

Blatchford, P., Bassett, P., Goldstein, H., & Martin, C. (2003). Are class size differences related to pupils' educational progress and classroom processes? Finding from the Institute of Education Class Size Study of children aged 5-7 Years. *British Educational Research Journal, 29*(5), 709–730, Special Issue 'In Praise of Educational Research', Guest Editors: S. Gorrard, C. Taylor & K. Roberts.

Blatchford, P., Kutnick, P., Baines, E., & Galton, M. (2003). Toward a social pedagogy of classroom group work. In P. Blatchford, & P. Kutnick (Eds.), *Special Edition of International Journal of educational research, Vol. 39* (pp. 153–172).

Blatchford, P., & Lai, K. C. (2010). Class size: arguments and evidence. In B. McGraw, E. Baker, & P. P. Peterson (Eds.), *International encyclopedia of education* (3rd edition). Elsevier.

Blatchford, P., Moriarty, V., Edmonds, S., & Martin, C. (2002). Relationships between class size and teaching: a multi-method analysis of English infant schools. *AmericanEducational Research Journal, 39*(1), 101–132.

Blatchford, P., & Mortimore, P. (1994). The issue of class size in schools: what can we learn from research? *Oxford Review of Education, 20*(4), 411–428.

Blatchford, P., Russell, A., & Brown, P. (2009). Teaching in large and small classes. In L. J. Saha, & A. G. Dworkin (Eds.), *International handbook of research on teachers and teaching* (pp. 779–790). New York: Springer, (ISBN: 978-0-387-73317-3).

Bourke, S. (1986). How smaller is better: some relationships between class size, teaching practices, and student achievement. *American Educational Research Journal, 23*(4), 558–571.

Bronfenbrenner, U. (1979). *The Ecology of human development*. Cambridge, MA: Harvard University Press.

Bruhwiler, C., & Blatchford, P. (2011). Effects of class size and adaptive teaching competency on classroom processes and academic outcome. *Learning and Instruction, 21*, 95–108.

Cooper, H. M. (1989). Does reducing student-to-teacher ratios affect achievement? *Educational Psychologist, 24*(1), 79–98.

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

Croll, P. (1986). *Systematic classroom observation*. London: Falmer.

Delamont, S., & Hamilton, D. (1986). Revisiting classroom research: a continuing cautionary tale. In M. Hammersley (Ed.), *Controversies in classroom research* (pp. 25–43). Milton Keynes, England: Open University Press.

Dunkin, M. J., & Biddle, B. (1974). *The study of teaching*. New York: Holt, Reinhart and Winston.

Ehrenberg, R. G., Brewer, D. J., Gamoran, A., & Willms, J. D. (2001). Class size and student achievement. *Psychological Science in the Public Interest, 2*(1), 1–30.

Evertson, C. M., & Randolph, C. H. (1989). Teaching practices and class size: a new look at an old issue. *Peabody Journal of Education, 67*(1), 85–105.

Finn, J. D., & Achilles, C. M. (1999). Tennessee's class size study: findings, implications, misconceptions. *Educational Evaluation and Policy Analysis, 21*(2), 97–109.

Finn, J. D., Pannozzo, G. M., & Achilles, C. M. (2003). The "why's" of class size: student behavior in small classes. *Review of Educational Research, 73*(3), 321–368.

Galton, M. (1998). Class size: a critical comment on the research. *International Journal of Educational Research, 29*, 809–818.

Glass, G., Cahen, L., Smith, M. L., & Filby, N. (1982). *School class size*. Beverley Hills, C.A: Sage.

Goldstein, H. (1995). *Multilevel statistical models*. London: Hodder Arnold.

Goldstein, H., Rasbash, J., Plewis, I., Draper, D., Browne, W., Yang, M., et al. (1998). *A user's guide to MLwiN*. London: Multilevel Models Project, Institute of Education, University of London.

Graue, E., Rauscher, E., & Sherfinski, M. (2008). *Using multiple data sources to understand the synergy of class size reduction & classroom practice in Wisconsin*. Paper to American Educational Research Association Annual Meeting, New York.

Grissmer, D. (1999). Class size effects: assessing the evidence, its policy implications, and future research agendas. *Educational Evaluation and Policy Analysis, 21*(2), 231–248.

Hanushek, E. A. (1999). Some findings from an independent investigation of the Tennessee STAR experiment and from other investigations of class size effects. *Educational Evaluation and Policy Analysis, 21*(2), 143–163.

Hattie, J. (2005). The paradox of reducing class size and improving learning outcomes. *International Journal of Educational Research, 43*, 387–425.

Konstantopoulos, S. (2008). Do small classes reduce the achievement gap between low and high achievers? Evidence from Project STAR. *The Elementary School Journal, 108*(4), 275–291.

Kounin, J. S., & Gump, P. V. (1974). Signal systems of lesson settings and the task-related behavior of pre-school children. *Journal of Educational Psychology, 66*(4), 554–562.

McIntyre, D., & Macleod, G. (1986). The characteristics and uses of systematic observation. In M. Hammersley (Ed.), *Controversies in classroom research*. Milton Keynes, England: Open University Press.

Mitchell, D. E., Beach, S. A., & Badarak, G. (1989). Modeling the relationship between achievement and class size: a re-analysis of the Tennesse Project STAR data. *Peabody Journal of Education, 67*(1), 34–74.

Molnar, A., Smith, P., Zahorik, J., Palmer, A., Halbach, A., & Ehrle, K. (1999). Evaluating the SAGE program: a pilot program in targeted pupil-teacher reduction in Wisconsin. *Educational Evaluation and Policy Analysis, 21*(2), 165–177.

NICHD Early Child Care Research Network. (2004). Does class size in first grade relate to children's academic and social performance or observed classroom processes. *Developmental Psychology, 5*, 651–664.

Pate-Bain, H., Achilles, C. M., Boyd-Zaharias, J., & McKenna, B. (1992). Class size makes a difference. *Phi Delta Kappan, 74*(3), 253–256.

Paterson, L., & Goldstein, H. (1991). New statistical methods for analyzing social structures: an introduction to multilevel models. *British Educational Research Journal, 17*(4), 387–393.

Prais, S. J. (1996). Class size and learning: the Tennessee experiment - what follows? *Oxford Review of Education, 22*, 399–414.

Shapson, S. M., Wright, E. N., Eason, G., & Fitzgerald, J. (1980). An experimental study of the effects of class size. *American Educational Research Journal, 17*, 144–152.

Slavin, R. E. (1989). Class size and student achievement: small effects of small classes. *Educational Psychologist, 24*(1), 99–110.

Wilson, V. (2006). *Does small really make a difference? An update. A review of the literature on the effects of class size on teaching practice and pupils' behaviour and attainment*. Scottish Council for Research in Education (SCRE) Centre: University of Glasgow.

Yodder, P., & Symons, F. (2010). *Observational measurement of behaviour.* New York: Springer.