Teachers’ and Pupils’ Behavior in Large and Small Classes: A Systematic Observation Study of Pupils Aged 10 and 11 Years

Peter Blatchford, Paul Bassett, and Penelope Brown
Institute of Education, University of London

The authors examined class size effects on teacher–pupil interactions, pupil engagement, and pupil–pupil interaction. They extended previous research by recognizing the hierarchical nature of observation data and the possible influence of other variables. The study used a time sampling method involving 257 children (aged 10–11 years) in 16 small (25 or under) and 26 large (31 and over) classes. In small classes, there were more individualized task-related contacts between teacher and pupils and a more active role for pupils. These results confirmed those from earlier research on children aged 4 and 5 years. Against expectation, class size did not affect pupil on-task behavior or peer interaction. There was a moderating role for school subject and a beneficial effect of teaching assistants.

Keywords: class size, teacher–pupil interactions, classroom engagement, peer interaction, pedagogy

Research on teaching and pedagogy has shown a lack of interest in classroom contextual influences. There has been an underlying assumption in many studies of a direct model, with teaching affecting pupils’ achievements and learning in a causal way (Blatchford, Kutnick, Beines, & Galton, 2003). Some academics and policymakers have also downplayed the importance of class size differences (Rivkin, Hanushek, & Kain, 2000). Yet teachers do not meet pupils out of context, and class size, or the number of children in the classroom, can be seen as a contextual influence on classroom life, one that plays a part in affecting the behavior of teachers and pupils. The conceptual roots of this view can be found in Bronfenbrenner (1979) and the ecological psychology approach. The basic idea is that the classroom context has distinctive forces or “signals,” different from other contexts, that 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. However, to date, little is known about specific ways in which these pressures might work.

There has been much international debate about the educational consequences of class size differences. In the United States, the debate has centered on the educational benefits and cost-effectiveness of class size reduction initiatives, whereas, in the United Kingdom, the debate has been more about the negative effects of large classes. There is some agreement, drawing on experimental (e.g., Finn & Achilles, 1999) and naturalistic studies (Blatchford, Bassett, Oldstein, & 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 widely appreciated that attention needs to be given to better understanding the classroom processes that might be related to class size differences (Anderson, 2000; Finn & Achilles, 1999; Finn & Pennozzo, & Achilles, 2003; Grissmer, 1999). Anderson (2000) has set out a comprehensive model of possible factors linking class size to student achievement; this model includes greater knowledge of students, more instructional time, greater student engagement, and more in-depth treatment of content in smaller classes. Finn et al. (2003) concluded on the basis of their review that the research evidence supports two main conclusions: Students in small classes in the elementary grades are more engaged in learning behaviors and display less disruptive behavior than students in larger classes. Finn et al. also concluded that effects on processes appear to fade out by later grades and that class size seems to affect student engagement more than teaching, although there is some evidence that teachers’ interpersonal styles benefit from small class reductions. A large-scale longitudinal study of class size and pupil–adult ratio differences in English primary schools (the CSPAR project), using a multivariate approach, resulted in a data-summary model of class size effects on classroom processes, with main areas being teaching and pupil behavior, including pupil attentiveness and peer relations, and within-class groupings (Blatchford, 2002a; Blatchford, Bassett, et al., 2003). Finn et al. pointed to the need for studies that make use of strong data, such as those obtained by systematic observations, to provide reliable evidence, and they identified the observational study by Blatchford (2003b) as one of the few studies of sufficient quality to be included in their review. This study, which was part of the CSPAR project just cited, found clear relationships between class size and teachers’ interactions with very young pupils in school, pupil classroom engagement, and pupil–pupil interactions. It was, however, restricted to very young children in their first year of English primary school (aged 4 to 5 years).

In the United Kingdom, the policy context for an interest in class size differences has changed recently. In response to concerns about large classes, the government introduced a maximum of 30
pupils to a class for children in the first 3 years of school (aged 5 to 7 years). However, this has caused concern about class sizes over the rest of the primary school stage, that is, for pupils aged 7 to 11 years (called Key Stage II), for which no limits of any sort have been introduced. In 2004, official government-collected data showed that about a quarter of pupils in England were in classes with more than 31 pupils. There are concerns that class sizes of this magnitude have a negative effect on teaching, pupil engagement in work, and pupils' relationships with each other. However, there have been no U.K. studies that have considered the effects of class size differences on teacher and pupil behavior at this stage. In this article, we seek to establish whether the effects of small class sizes found in the earlier study (Blatchford, 2003b), in terms of increased individual attention, decreased levels of off-task behavior, and increased rates of pupil-pupil interaction, would still be evident at the end of the primary school stage (i.e., pupils aged 11 years).

We also in this article extend our earlier observation study (Blatchford, 2003b) and the literature on class size effects on classroom processes in two main ways. First, we examine the extent to which class size effects are affected by school subject, and second, we examine in a systematic way the effect of extra adults in the classroom on teacher and pupil behavior. More details on these two topics are given below. The study also contributes by using a more sophisticated form of data analysis than the previous study to model the hierarchical, clustered nature of observation data and the possibly confounding influence of other variables. Three aspects of teacher and pupil behavior are considered in this study: teacher-pupil interactions, pupil engagement and off-task behavior, and peer relations in class.

Teacher-Pupil Interactions

Class size can be positioned as one context factor in Dunkin and Biddle's (1974) early seminal model of effects on teaching, but in general, there has been little systematic research on how class size differences affect teaching behavior. On logical and common sense grounds, it seems likely that the number of children in a class affects the amount of time that teachers spend in procedural matters inversely affects the amount of time that can be spent on instruction and in dealing with individual children. This expectation is consistent with teachers' views (Beckett, 1996; Pate-Bain, Achilles, Boyd-Zaharias, & McKenna, 1992) and some previous research (Cooper, 1989; Glass, Cohen, Smith, & Felty, 1982). Achilles (1999) found, in a systematic observation study of two schools matched on background factors, that teachers in small kindergarten through Grade 2 classes (about 14 students) engaged in more on-task behavior over the year, whereas teachers in large classes (about 24 students) engaged in more off-task behavior over the year. However, Sheppson, Wright, Easton, and Fitzgerald (1980) found that results from systematic observations conflict with the teacher's own views; 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. There was, therefore, a gap between the teacher's own view and classroom process and observational research findings. Several studies have agreed that the most important classroom process affected by reduced pupil-adult ratios and class size is individualization of teaching (Betts & Shkolnik, 1999; Molnar et al., 1999). A recent multitrait study by the NICHD Early Child Care Research Network (2004) summarized its collected data by arguing that smaller first-grade classrooms appear more child centered than larger classes and that larger classes become more structured. However, the curriculum for older pupils aged 10 to 11 years, at least in the United Kingdom, is itself much more structured, and the context for class size effects may therefore be different.

Results from the earlier CSPAR study of 4- to 5-year-old pupils (Blatchford, 2003b) showed that, in small classes, children were more likely to interact with their teachers, more one-to-one teaching took place, children were more often the focus of a teacher's attention, more teaching took place overall, and children more often attended to their teachers. These results show that individual children in small classes received more interactions with their teachers of a task-related nature. The trend toward individualization in small classes does not seem indicative of a passive role for children; children in large classes spent less time actively interacting, either by responding or initiating contacts with the teacher. There was more interaction about classroom procedures (e.g., collecting materials, taking the register) in large classes. Overall, it was proposed that there was, in smaller classes, more likelihood of teacher support for learning (Blatchford, Moriarty, Edmonds, & Martin, 2002). In the present study, the aim was to establish whether these effects of small classes on teacher-pupil interactions would still be evident when pupils were 10 to 11 years of age.

It might be argued that one solution to the teacher's difficulties in contacting children in large classes would be to alter his or her approach so that there would be more teaching to larger groups or to the whole class. There is no evidence from the CSPAR study (Blatchford, 2003b) that teaching to the whole class increased in larger classes, but this result might have owed much to the perceived inappropriateness of whole-class teaching with such young children. We expected teachers with larger Year 6 classes to make more use of whole-class teaching, whereas teachers with smaller classes would make more use of individual settings and group teaching. We also wanted to see whether the more standardized curriculum and whole-class teaching approaches common with pupils aged 10 to 11 years had affected the tendency for more individual attention and active interactions in smaller classes. Given the extra demands on teachers' time, it was expected that there would be more waiting for the teacher in large classes, for example, to answer a query or provide materials.

Pupil Engagement and Off-Task Behavior

Children's successful adjustment to school is likely to involve their attentiveness in class, for example, as reflected in the extent of their attention to the teacher and to work. Many studies have supported the view that key aspects of educational success are engagement and active learning time, time on task, or some equivalent term (e.g., Creemers, 1994). There is a good deal of evidence that these are major variables that affect pupils' achievement (Rowe, 1993).

Common sense and logic suggest that with more children in the class, there is more potential for distraction and more possibility of being off task. Conversely, in small classes, there are more opportunities to engage children and keep them on task. Finn and Achilles (1999) have argued for a connection between small
classes and increased student engagement in learning. Cooper (1989), in his review, found studies to support this view, but there are limitations in previous research because of the often informal designs and reliance on teacher reports (Finn et al., 2003). There are also limitations in the conceptualization of off-task behavior in that the generic term may confine separate forms of off-task behavior, which may in turn be connected differently to class size differences. Off-task behavior may have two main forms: an externalizing form, in the sense of overtly disruptive behaviors and mucking about, or a more internalizing form, in the sense of disengagement and distracted from work. In the earlier CSPAR study (Blatchford, 2003b), we found, in the case of 4- to 5-year-old pupils, more off-task behavior in larger classes but especially more passive off-task behavior—more disengagement—when pupils were working on their own. We predicted that there would continue to be, for older pupils, more off-task behavior of both types in large classes, but because teachers are likely to sanction most obviously against more overt, externalizing forms, we predicted that the clearest effects of class size would again be on passive forms of off-task behavior. We also predicted that there would be more child procedural behavior in large classes (e.g., in terms of talk about and collection of materials, taking the register).

Peer Relations in Class

There is a lot of evidence that children’s early social behavior toward peers is an important predictor of later social and personal adjustment (Parker & Asher, 1987; Rubin, Bukowski, & Parker, 1998). There is also a large literature on the value of collaborative or cooperative group work in classrooms (Galton, 1990) and naturalistic studies of children’s interactions in classrooms, which include those with peers (Baines, Blatchford, & Kuttick, 2003; Galton, Simon, & Croll, 1980). However, as in the case of teacher–pupil interactions and pupil–task behavior, there has been little research on the effects of school contextual classroom factors such as class size on peer relations.

It might be expected that, in larger classes, the increased pressure for space and difficulties of getting individual attention from the teacher would increase the likelihood of interaction with peers and that there would be more negative and aggressive behaviors between children. This is supported by research on children at nursery level (A. B. Smith, McMillan, Kennedy, & Ratcliffe, 1988; P. K. Smith & Connolly, 1980), but other research with older pupils seems less clear. Shapson et al. (1980) found no difference in conflicts between pupils.

Results from the CSPAR reception-year study (Blatchford, 2003b) showed that children in larger classes spent more time with each other; this applied to all child–child categories, including interactions about work, interactions about social matters, and also mucking about. It was argued that pupils could compensate for reduced levels of interaction with teachers in large classes by interacting more with each other, and this might go some way to explaining why there was no support for the expectation that relations between children would be worse in large classes (Blatchford, Edmonds, & Martin, 2003). Given the increasing importance of the peer group as children progress through middle childhood, it was predicted that the increase in peer interaction with larger classes would be more apparent in older children, aged 11 years. As stated above, this article extends previous research by also examining the effect of the presence of extra adults and the role of school subject.

Presence of Adults in the Class

There is recognition that measures of class size and pupil–adult ratios are different and need to be clearly identified in research (Blatchford, Goldstein, & Mortimore, 1998; Ehrenberg, Brewer, Gamoran, & Williams, 2001; Finn & Achilles, 1999). A class size of 15 may not be equivalent to a class size of 30 with two teachers or adults because the social psychological processes at work can be different; Finn et al. (2003) argued that there are distinct qualities of a small class that are lost when there are more children in a classroom, despite the presence of extra adults. There is also a problem that overall measures of the number of pupils (taken from the class register) and of the number of adults may not necessarily adequately capture the number and type of adults in a class at any given moment. In England, at least, there can be a fair amount of movement in and out of classrooms by adults and pupils. For this reason, it is important to be able to record the presence of teachers and adults in the classroom on a moment-by-moment basis.

There has been a recent U.K. government drive to increase the numbers of support staff in schools. The generic term used for paid support staff working in classrooms is teaching assistants (TA), and there has been a large increase in their numbers. Studies that have assessed the effects of TAs on pupil academic outcomes have tended not to find positive effects (Blatchford, Russell, Bassett, Brown, & Martin, 2004; Finn, Gerber, Farber, & Achilles, 2000; Muijs & Reynolds, 2000). However, one can conceive of the impact of TAs not just in terms of pupil learning and attainment but also in relation to pupil and teacher behavior. Once again, there is only relatively anecdotal evidence, and in this study, we wanted to measure in a systematic way whether the presence of TAs had an effect on interactions involving pupils and teachers in the same classrooms, as well as the extent of classroom engagement and attention to the teacher. We also wanted to see whether any influence of class size interacted with the presence of adults, for example, whether the presence of a TA had a more marked effect in larger classes compared with small.

School Subject

There has been little research on the possible moderating effect of school subject on class size influences on classroom processes. Rice (1999), on the basis of a secondary analysis of teacher survey data, found that in math, as class size increased, less time was spent on small groups and individuals, innovative instructional practices, and whole-group discussions, although increases in class sizes beyond 20 had little effect. There were no relationships between class size and instructional time allocation measures in science. In the earlier CSPAR study (Blatchford, 2003b), with the youngest children in school, teaching methods and parts of the school day were not clearly delineated in terms of subject differences. However, by the end of the primary stage, when pupils are aged 10 to 11 years, there is a much clearer separation between subjects and a clearer apportioning of prescribed time, in terms of discrete lessons, to the three core subjects of math, English, and science. Different teaching approaches, appropriate for different subjects, have also developed by this time. Traditionally, some
subjects, such as science and English, have been taught using more discussion or practically based groups, whereas other subjects, such as math, have relied on individual workbooks or traditional teacher-to-whole-class expositions. These differences in teaching approaches might be expected to be significant in any relationships between the number of pupils and teacher and pupil behavior. In England, national literacy and numeracy strategies have been introduced, with accompanying suggestions for teaching strategies and end-of-year assessment arrangements. This has raised concerns that other areas of the curriculum, such as history, geography, and art, might be affected. We wished to see whether there were any differences between school subjects in moderating any class size effects found.

Method of Data Collection: Systematic Classroom Observation

One way to measure more reliably the extent to which class size differences are related to the type and nature of teacher-pupil interactions and pupil work-related behavior is through the use of systematic observation. This collects data in terms of categories previously defined on conceptual grounds and refined in pilot 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 for which data are needed on relatively easily observed, high-frequency behaviors (Croll, 1986; McIntyre & Maceod, 1986). This was the method used in the Shapson et al. (1980) study that, although widely cited, is dated now.

There are difficulties, but also potential advantages, in the use of systematic observation, particularly with regard to its ability to capture the co-occurrence of behaviors at particular moments in time. The difficulties include the time-consuming nature of data collection and processing, which sometimes leads researchers to use total frequencies of behaviors as measures in analysis. This choice limits the ability to study relationships between variables and the ability to examine effects of class size on a given observation category, say, teacher individual attention, controlling for the occurrence of another category, say, the presence of a TA or whether math or English is occurring, at the same moment in time. One advance of this article over the earlier study (Blatchford, 2003b) is that the current article measures the effects of other possible factors connected to, or that moderate, the effect of class size differences on classroom behavior.

There are also particular difficulties with observation data connected to the clustering of behavior in time and within classrooms and schools. The treatment of these difficulties requires a more sophisticated statistical analysis than has commonly been the case in research making use of systematic observation data. Multilevel statistical models are required, as the observations are not always independent of each other. More details are given in the Method section, below.

Sample of Pupils

Teachers were asked to select nine pupils, three from each ability range—low, medium, and high. Six of these target children were then chosen by the researcher, two from each ability band, one girl and one boy. If a child was absent for more than a day, he or she was replaced by a reserve drawn from the nine. In some cases, there were observations on
more than six children (e.g., the originally selected child and the reserve), and there were 237 children in all, 128 girls and 129 boys, 83 low ability, 87 medium ability, and 87 high ability.

Organization of Observations

The basic principle was to observe when classroom-based activities could have taken place. Observations were not conducted during parts of the day when all the pupils went out of the classroom. The aim was to observe each child over 2 days. Time available for observation could vary somewhat from day to day. Observations were conducted in blocks of 10-s time intervals, with gaps of 10 s between observations to allow recording of what took place in the previous 10 s. After each block of 10 observations, attention switched to the next pupil on the list. There were 22,312 observations in total, with an average of 87 observations per child.

Observation Categories

The schedule was similar to that used in the earlier study of 4- and 5-year-olds (Blatchford, 2003b) and comprised categories that provided a description of time spent in different work settings (individual, pupil group, teacher-led group, whole class), a description of time spent in different school subject areas (English, mathematics, science, and other, e.g., history, geography, religious education (RE)), and a description of how children behaved when in three social modes (with their 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 and other children were observed only when they came into contact with the target child. The schedule used a form of predominant activity sampling with those behaviors selected within sets of behaviors (e.g., social modes) that occurred for the longest period within the 10-s interval. To examine the effect of class size, we chose selected categories on conceptual grounds and on the basis of relatively high frequency of occurrence. Brief definitions of these categories follow.

School Subject

At the time of the observation, the child was engaged in one of the following: mathematics, English, science, or other subjects (which would include geography, history, art, RE, music, technology).

Work Context

Individual context: The child was working on his or her own; the work was not group based (though the child could be seated in a group) or teacher led. Group context: The child was in a group working together but not led by the teacher. Whole-class context: teacher-led whole-class settings in which the target child was involved.

Teacher-Pupil Interaction

Teacher-to-child talk. Adult teach: adult behavior directly concerned with the substantive content of subject knowledge, that is, communicating concepts, facts, or ideas by explaining, informing, demonstrating, questioning, suggesting. Adult on task: same as adult teach plus contexts concerning the organization and preparation of children's task activities but not their substantive content.

Child audience versus focus. Child is focus: Target child was the focus of the teacher's attention, and this could have been in the context of one-to-one, group, or whole-class sessions, for example, the target was asked a question about addition in the course of a lesson in which the teacher was addressing the whole class. These were coded separately as "short," that is, not for the whole 10-s interval, and "long," that is, contact continued through the whole 10-s period—for example, a question from the teacher was followed by an answer from the child and a further probe or comment from the teacher. This therefore gave some measure of extended or sustained interactions between child and teacher. Child is audience. Another child was the focus of the teacher's attention in the group or class, involving the target child, or the teacher interacted to the same extent with all children.

Child-to-Teacher Interaction

Child-to-teacher attend/listen. The child simply listened to the teacher during the interval and did not interact by responding or initiating.

Child on task to teacher. This covered all child behaviors in contact with teacher that were concerned with work.

Child off task to teacher. This covered child behavior when in contact with the teacher that was obviously inappropriate or unrelated to the situation (e.g., not attending).

Wait for interaction with the teacher. The target waited for the teacher.

Pupil-Pupil Interactions

Target and child on task. This covered all contacts with other children that were concerned with work and allocated tasks.

Target and child off task. This covered behavior with other children that was deliberately off task; it included nuzzling about and fooling around, as well as times the target child was aggressive (verbally or physically) toward another child or other children. It did not include times when children spoke about nonwork activities if this was not deemed unacceptable by the teacher (this would have been coded "social").

Individual Behavior/Not Interacting

Individual on task. Target child was involved in his or her own work activity.

Individual off task (active). Target child focused on something other than the task at hand.

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

Computed Categories

Child on task. This category comprised total on-task behaviors, that is, behaviors related to the substantive nature of allocated work or preparation for the work across the three social modes, that is, child-to-teacher on task, target and child on task, and individual on task.

Child off task. This category comprised total off-task behaviors, that is, all off-task behaviors in the three social modes, that is, child-to-teacher off task, target-to-child off task, and individual off task (active and passive).

Child procedure. This category comprised total child procedure behaviors, that is, all target behaviors related to classroom management and organization of classroom routine in the three social modes, that is, child-to-adult procedure/routine, target-to-child procedure/routine, and individual procedure/routine.

Active interaction with teacher. This category comprised the sum of the three child-to-teacher categories in which the child's role was an active and not a passive (i.e., attend/listens) one, that is, the child initiated, responded, or sustained interactions with the teacher.

Any target and child interaction. This category comprised the sum of all the pupil-pupil categories, that is, all task, social, procedure, and off-task behaviors in contact with other children.

Other Information

At the time of each observation (i.e., each 10-s time interval), a note was made of the number of children in the classroom; the term class size
therefore describes the number of pupils in the classroom at the time, whether in class or not (what we have called the experienced class size), rather than a notional number on the register. For each observation, a note was also made of any extra adults in the classroom; extra adults refers in almost all cases to paid TAs. A note was also made of whether the observations took place when the class was setted or the normal registration class and, further, whether the set comprised low-, medium-, or high-ability pupils. The sex and ability level of target pupils were also noted.

Observers. There were four observers. They were all experienced researchers who were familiar with working in schools and 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. All observers 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 at least a day's observation in a class not involved in the study and then a follow-up training session to discuss field visits and iron out difficulties.

Reliability checks. Reliability coefficients for the main sets of mutually exclusive categories were high; setting; subject; teacher–child social setting; child role, teacher content; and child-to-teacher child contribution; child content, and not interacting all had reliability coefficients kappa greater than 0.80. Kappa for child–child content was 0.77.

Statistical methods and analysis. As described above, a feature of the analysis of the observation data was the way that the analysis was conducted with the 10-s observation interval as the unit of analysis. This allowed a greater accuracy and flexibility than simple, more commonly used, total frequencies of behaviors for each pupil. The observation variables took the form of binary variables, in the sense of each either being performed or not being performed during one time interval.

A further feature of this observation study is that there is a hierarchical structure to the data, which is summarized in Figure 1. A consequence of the hierarchical structure is that it was likely that observations from pupils in the same class would be more similar than observations from pupils in different classes. Similarly, observations from the same pupil were more likely to be similar than observations from different pupils. Therefore, the observations could not be regarded as independent of each other, and so, multilevel statistical methods (Goldstein, 1995) were required to allow for the non-independence of the data. Failure to allow for the structure of data can lead to misleading parameter estimates and also to an exaggeration of the significance of results (Paterson & Goldstein, 1991).

Four-level multilevel models were used for the analysis, with the levels described in Figure 1. As all outcome variables were binary in nature, logistic regression models were used to examine the effects of the various explanatory variables on the outcomes. Let \( y_{ij} \) be the observed binary response \((0, 1)\) of observation \(i\) from the group of observations \(j\), from pupil \(k\) in class \(l\), and let the probability of a particular outcome being observed be denoted by \( \pi_{ijk} \), where \( \pi_{ijk} = \Pr(y_{ij} = 1) \). The basic form of the regression equation used in the analysis for a single explanatory variable \( x \) is given by

\[
\logit(\pi_{ijk}) = \beta_0 + \beta_1 \cdot x_{ij} + u_{1j} + u_{2k} + u_{3l},
\]

where \( u_{1j} \) is the random effect at the class level, \( u_{2k} \) is the random effect at the pupil level, and \( u_{3l} \) is the random effect the observation group level. All random effects are assumed to be normally distributed as follows:

\[
u_{1j} \sim N(0, \Omega_{1j}), \quad u_{2k} \sim N(0, \Omega_{2k}), \quad \text{and} \quad u_{3l} \sim N(0, \Omega_{3l}).\]

The regression models were fitted using the MLwiN software package (Goldstein et al., 1998).

The exception to this basic form of the regression model was for the work contexts (individual, group, whole class). Within each group of observations, the pupils were always performing the same type of work. Therefore, data for these variables were analyzed at the group level with one observation per group, and hence, three-level multilevel models were used.

The initial set of analyses examined the differences between small and large classes without any adjustments for possible confounding factors. Subsequently, the effects of class size and several other explanatory factors thought likely to influence the observation outcomes were examined jointly. The aim was to estimate the effect of class size difference, controlling for the effects of the other explanatory factors. A summary of the other measures included in the analysis is given in Table 1. A backward selection approach was performed to retain only the significant variables for each outcome, although class size was left in all models, as the effect of this variable was the main focus of the analysis. Possible interactions between class size and the other factors in Table 1 were also examined to see if the effect of class size on the outcome varied for different levels of the other factors. Where significant interactions have been found, the results are reported separately for each subgroup. Main effects for the presence of TAs are included in tables below when they are statistically significant.

Results

The results of the statistical analyses examining the effects of the number of pupils in the class on each of the main observation categories are summarized in Table 2. The summary statistics reported are the effect sizes of the odds of each outcome occurring for a child in a large class relative to a child in a small class. No other possible explanatory variables other than class size are considered at this stage. An odds ratio above 1 suggests the outcome would be more likely in a large class, whereas an odds ratio below 1 suggests that the outcome would be more likely in a small class. Also given are 95% confidence intervals for each odds ratio and a p value indicating the significance of each result.

There are some clear differences between small and large classes. Two-related behaviors were more common in large classes: child-to-teacher listen/attend and child is audience. The first cat-

Figure 1. Structure of the data.
Table 1

<table>
<thead>
<tr>
<th>Variable</th>
<th>Reference category</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of pupils</td>
<td>Small numbers (25 or fewer)</td>
<td>Effect of large numbers in class (i.e., 30 or more) relative to small numbers</td>
</tr>
<tr>
<td>Extra adults/TAs Sets</td>
<td>No extra adults/TAs Classes</td>
<td>Effect of extra adults/TAs relative to none</td>
</tr>
<tr>
<td>Ability</td>
<td>Medium ability</td>
<td>Effect of low-ability or medium-high ability sets relative to classes</td>
</tr>
<tr>
<td>Sex</td>
<td>Girls</td>
<td>Effect of low- or high-ability pupils relative to medium-ability pupils</td>
</tr>
<tr>
<td>Subject</td>
<td>Math</td>
<td>Effect of boys relative to girls</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Effect of each subject compared with math</td>
</tr>
</tbody>
</table>

Note. TAs = teaching assistants.

egory denotes times when the child’s contribution to interactions with the teacher was passive; the child was simply listening to the teacher. Child audience refers to times when the child was not the focus of the teacher, that is, the child was not singled out by the teacher either on a one-to-one basis or in a group or whole-class situation. Both therefore describe a passive role in contact with the teacher, and this was more likely in larger classes.

Conversely, in smaller classes, pupils had a more active role in contact with the teacher. This can be seen in the greater likelihood of active forms of behavior in contact with the teacher, that is, initiating and responding to the teacher and sustaining contact with the teacher.

There was also a greater likelihood in a small class of the child being the focus of a teacher’s attention. This is evident in terms of both short and long focus, as well as the two added together.

It can also be seen that pupils experienced more teaching, in the strict sense of contacts involving the substantive content of a subject (i.e., adult teach), in small classes compared with large. It is interesting to note that, although teaching was more common in small classes, there is slight evidence that times when the teacher was addressing the whole class (which could cover all types of content, including procedure/routine) were more likely in large classes.

Also of interest are the results that show no differences between large and small classes. Against expectation, there are no differences in the on-task and off-task categories, procedural behavior, and waiting for the teacher.

The results in Table 2 consider class size as the only explanatory variable in relation to observation variables. We now examine effects of class size while taking into account other possible explanatory factors. The effect of the presence of adults is also examined as a main effect because of the interest in this variable. Possible interactions between class size and the other factors in Table 1 were examined to see if the effect of class size on the outcome varied for different levels of the other factor. Where significant interactions have been found, we report the effects separately for each subgroup.

Work Context

Results concerning effects of class size and adults on the three work-context categories are shown in Table 3. There was no sign that class size affected the amount of time pupils worked individually. The effect of the number of pupils on the other two work contexts varied by subject. Groups were more common for larger numbers of pupils for science but conversely less common with large numbers of pupils for other subjects. Whole-class contexts were more common for other subjects when there were a large
number of pupils compared with a small number, but there were no differences for the three remaining subjects.

The effect of the presence of TAs on each of the three work contexts varied by subject. There was more individual context for math with an adult present, but no difference for the other three subjects. There was more group context for English with an extra adult present, but less group context for other subjects with the presence of an adult. There was less whole-class context with an adult present for math and English, but no difference for science or other subjects.

**Teacher–Pupil Interactions**

Results for adult teach and adult on task are presented in Table 4. For both categories of teacher talk to pupils, the effect of the number of pupils varied by subject. There was less adult teach and adult on task in large compared with small classes in math. No differences were found for the three other subjects.

For adult teach, there was an effect of extra adults in the classroom. This varied by subject; the presence of a TA increased adult teach in science but decreased it in English.

**Child Audience and Focus**

The connections between class size and child audience and child focus are shown in Table 5. Child audience refers to times when the teacher addressed another pupil or all pupils together. A large number of pupils resulted in more child audience in math, science, and other subjects, but there was no difference between numbers of pupils for English. There was a significant effect of having a TA in the classroom, with the odds of child audience being halved when a TA was present.
Table 5

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Explanatory variable</th>
<th>Subgroup</th>
<th>OR (95% CI)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child audience</td>
<td>Number of pupils</td>
<td>Math</td>
<td>3.10 (1.94, 4.95)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>English</td>
<td>1.13 (0.75, 1.70)</td>
<td>.55</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Science</td>
<td>3.00 (1.53, 5.87)</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Other</td>
<td>2.62 (1.54, 4.47)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Child focus (short and long)</td>
<td>TAs/extra adults</td>
<td>—</td>
<td>0.54 (0.41, 0.70)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Number of pupils</td>
<td>Math</td>
<td>0.29 (0.16, 0.50)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>English</td>
<td>0.89 (0.48, 1.63)</td>
<td>.62</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Science</td>
<td>0.15 (0.07, 0.35)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Other</td>
<td>0.25 (0.18, 0.38)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>TAs/extra adults</td>
<td>—</td>
<td>1.68 (1.24, 2.27)</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

Note. OR = odds ratio; CI = confidence interval; TAs = teaching assistants.

The category focus was coded whenever the target pupil was the subject of a teacher’s attention. There was more of this individual attention from teachers in small classes for math, science, and other subjects, but no difference for English. There was a noticeable effect of the presence of a TA in the classroom for total focus, with the odds being over 50% higher when a TA was present.

The results for focus short and focus long, when analyzed separately (not presented here), were less clear. They again pointed to less individualized attention with larger numbers of children, though this was affected by both subject and whether in classes or sets.

Child-to-Teacher Interactions

The results for child-to-teacher interactions are seen in Table 6. A small class size and presence of a TA increased active interaction with the teacher (i.e., the total of all initiations, responses, and sustained interactions with the teacher).

The number of pupils did not influence waiting for interaction with the teacher. Neither did the presence of a TA.

The effect of the number of pupils on child-to-teacher attend/listen was found to vary by set type. There was more attend/listen in large classes than in small classes, but no significant results for either of the two types of sets.

Pupil-to-Pupil Interactions

Table 7 shows connections between class size, pupil-pupil on- and off-task behavior, and all pupil-pupil activity. Class size was not related to pupil-pupil on- or off-task behavior but did play some role for the total of all pupil-pupil interactions. In a large class, pupil-pupil interaction decreased in math and other subjects, but there was no difference between large and small classes for English and science.

The presence of a TA in the classroom influenced pupil-pupil on-task behavior, with the effect found to vary by subject. There was more pupil-pupil on-task behavior when there was a TA present in English, but no evidence of an effect for the remaining subjects. There was no effect of TA presence on pupil-pupil off-task behavior. There was less total pupil-pupil interaction with a TA present for other subjects, but no evidence of an effect for math, English, or science.

Not Interacting Behavior

Table 8 shows factors influencing individual on- and off-task behavior. It also shows that the number of pupils did not have an effect on individual on- or off-task measures.

There were signs that the presence of a TA increased the occurrences of individual on-task behavior and decreased the occurrence of off-task behavior. No effect was found when active and passive off-task behaviors were examined separately, but this could have been due to the reduced number of occurrences of each separate behavior type.

Total On- and Off-Task Behavior

Table 9 shows the results for pupil on- and off-task behavior summed across the three different social modes. The results show
that there was no effect of the number of pupils on total child on-task behavior.

The effect of TAs on on-task behavior varied by subject. There was more on-task behavior when there were extra adults for other subjects, but no effect for math, English, or science.

There was less total off-task behavior with extra adults present compared with when extra adults were not present. The effect of number of pupils on off-task behavior was found to vary by subject, with more off-task behavior for large numbers compared with small numbers for other subjects, but no evidence of a difference for math, English, or science.

We next asked if there were any other factors connected to whether a pupil was in a small or a large class that might have affected the observation results. The most obvious possibility is that the allocation of pupils to classes is nonrandom, for example, that children with more difficulties at school are more likely to be found in small classes. To assess this possibility, we compared small and large classes and found that pupils in small classes did indeed have lower attainment scores compared with pupils in large classes. In addition, small classes contained a greater proportion of pupils eligible for free school meals and with special needs. Analysis of relations with the observation variables, however, showed that these three variables, that is, pupil attainment, free school meals, and special needs status, did not influence the observation outcomes. We can conclude, therefore, that results comparing small and large classes were not affected by any differences in pupil characteristics between class sizes.

Discussion

It is acknowledged that there are difficulties in drawing causal inferences from data in this study. The logic of the naturalistic design makes it impossible to be absolutely sure about causal direction. However, we have sought to control for possibly confounding and interacting variables when examining associations between class size and classroom behavior, and the results of the analyses suggest that, even after allowing for the effects of other variables, there were significant effects of both the number of pupils in the classroom and the presence of extra adults or TAs in the classroom on some of the observation categories. With many observation categories, however, there was a tendency for main effects of class size to be moderated by the school subject within which observations took place.

Work Context

There were three work-context categories—individual activity, groups, and whole class. There was no evidence that the amount of individual work was affected by class size. It was found that teachers in larger classes tended to use more groups, but for science only. Teachers do not therefore seem to have organized the pupils in terms of more work in groups in response to having large classes in the two main subjects of math and English. It is not possible to give a precise explanation for this, but there are two main possibilities: First, there is evidence that teachers’ pedagog-
Table 9
Class Size and Total On- and Off-Task Behavior

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Explanatory variable</th>
<th>Subgroup</th>
<th>OR (95% CI)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total child on task</td>
<td>Number of pupils</td>
<td></td>
<td>0.93 (0.74, 1.16)</td>
<td>.52</td>
</tr>
<tr>
<td></td>
<td>Task/extra adults</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total child off task</td>
<td>Number of pupils</td>
<td></td>
<td>1.25 (1.05, 1.50)</td>
<td>.11</td>
</tr>
<tr>
<td></td>
<td>Task/extra adults</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. OR = odds ratio; CI = confidence interval; TAs = teaching assistants.

When applied to these subjects do not recognize the value of group work (Blatchford, Kutznick, et al., 2003), and second, it may also be that the curriculum in these two subjects was too constrained in terms of content and reliance on whole-class teaching to allow for more work in groups. Moreover, science teaching in the United Kingdom has traditionally been more likely to involve work in groups, for example, to conduct experiments and practical tasks, and teachers with large classes may have been more likely to rely on this approach. This raises questions about whether teachers with large classes could make more use of groups in math and English. We return to this point below.

As expected, there was more whole-class teaching in larger classes, but this was only for the other subjects, such as geography, history, music, and art activities. One explanation for why teachers with larger classes did not rely on more whole-class teaching in the three main core subjects may be that constraints of the curriculum result in teaching methods that are less flexible and, therefore, less affected by the number of pupils. Teachers do, however, have more flexibility about the teaching and coverage of other non-core subjects in the curriculum, and teaching may therefore be more likely to be adapted when teachers are faced with large numbers of pupils. There are already concerns that the stress on the three core subjects of math, English, and science is tending to squeeze time spent in other areas of the curriculum. It looks as though teachers in larger classes are also adapting their teaching in other subjects to use more whole-class sessions. Other evidence from teachers themselves suggests that they do not always feel whole-class teaching is the best way to teach these subjects (Blatchford, Russell, Bassett, Brown, & Martin, 2003).

Pupil Is Individual Versus One of the Crowd

In these Year 6 classrooms, there was little individualized, or one-to-one, contact between teachers and pupils. Overall, there was a heavy reliance on whole-class teaching and individual work. Yet the results are clear in showing that an effect of class size on individual attention did take place. In small classes, there was more chance that pupils would be the focus of a teacher’s attention. Conversely, in a large class, there was more chance that a pupil would be in an audience mode, that is, listening to the teacher address all pupils equally or another pupil. Small classes therefore seem to allow more individual attention, whereas, in large classes, children are more likely to be one of the crowd. This reinforces results from the earlier reception study (Blatchford, 2003b) and extends results to pupils at the end of the primary stage (10 to 11 years).

There are indications that this main effect on individual attention was modified in some ways by school subject. There was significantly more child as audience when there were a large number of pupils for math, science, and other, but no difference between class sizes for English. In addition, there was significantly less adult focus when there were a large number of pupils in the same three subjects, but again, there were no differences between sizes for English. This suggests that the general effects of class size on individualized attention just cited are found in all subjects but English. The impression gained by observers in this study was that this is because English is the most discussion-based subject; even in teacher-to-whole-class sessions, the teacher tends to set up individual work by discussing and going through a topic with all in the class, and the dialogue is less likely to rely on questioning of individual pupils.

Active Versus Passive Interaction With the Teacher

For the most part, pupils, when in contact with their teachers, tended to be in the passive mode of listening to them. When considering active interaction that did take place, however, there was a greater likelihood of it in small classes. Pupils in small classes were more likely to initiate interactions, respond to the teacher, or sustain dialogue with the teacher into the next time interval. This result is consistent with results on much younger children. Conversely, in larger classes, there was more listening and attending to the teacher. This behavior is in a sense on-task behavior but is a passive form of it. It seems that this trend is most obvious in classes rather than sets (i.e., when classes were recognized by same ability). The number of pupils was not found to influence the occurrences of waiting for interaction with the teacher.

Amount of Teaching

Class size also affected the overall amount of teaching. In smaller classes, there was more teacher-to-pupil talk directly concerned with the substantive content of subject knowledge; communicating concepts, facts, or ideas; and so forth (adult teach).
This is in line with Achilles' (1999) results. However, for both types of teacher talk (adult teach and adult on task), the effect of the number of pupils varied by subject. There was significantly more teaching overall in small compared with large classes in math, but no differences for the other three subjects. It is not possible to state exactly the reasons for this finding, but observers' impressions were that teachers felt more able to extend their introductory explanations and demonstrations of procedures in small classes because they knew that it would take less time to move around and monitor individual work. Such introductions are more typical of math compared with other subjects. In other words, teachers in large and small classes may have balanced whole-class introductions and support for individual work differently.

Student Engagement in Class: On- and Off-Task Behavior

The results are clear in showing that there were no main effects of class size difference on any of the individual or total on- or off-task categories. Pupils in small and large classes spent the same amount of time on and off task. This was against expectation and contrary to results from the earlier CSPAR study of children aged 4 to 5 years (Blatchford, 2003b). The only qualification to this overall trend is that there were signs of more total off-task behavior in larger classes in other subjects (i.e., not maths, English, or science). One explanation of this finding may be that the prescribed curriculum in these three subjects meant that there was little opportunity for work-related behavior by pupils to vary according to the number of pupils.

Pupil–Pupil Interaction

The number of pupils in the classroom did not influence the occurrence of pupil–pupil on- and off-task behavior but did affect the overall amount of pupil–pupil interaction. This effect was found to vary by subject, with evidence that, in small classes, there was more interaction between pupils in math and other subjects but no difference between numbers for English and science. What is clear is that, against expectation, there was no evidence that there was more interaction between pupils in large classes. This is another result that differs from that found in the earlier CSPAR study of 4- and 5-year-olds (Blatchford, 2003b). It is likely that this result, as with other results, owes much to the strong pressure on pupils to accord to particular curriculum and work demands, which left little time for interactions with peers.

Effect of Extra Adults

We have shown that the extra adult in these classrooms was almost always a paid TA. One clear finding to emerge from the study is the beneficial effect of having a TA in the classroom. To date, the evidence about the impact of support staff has tended to be in terms of the limited effects they have on student academic outcomes. The present observation study, however, found that their presence in classrooms had effects not on interactions with pupils in which they were involved themselves but rather on interactions between the teacher and pupils in the same classroom.

The effect of TAs in the classroom on teacher and pupil behavior was for the most part independent of a class size effect. The clearest results concerned the effect of a TA's presence on individual attention by the teacher. With a TA in the classroom, pupils were more likely to be the focus of a teacher's attention and less likely to be in an audience mode. So, it seems that both small classes and the presence of a TA increase the possibility of individualized attention for pupils and that these two factors are independent of each other. An allied effect of the presence of TAs was in terms of increasing occurrences of pupils' active interaction with the teacher.

The presence of a TA was also found to significantly increase the occurrence of individual on-task behavior and significantly decrease off-task behavior. No effect was found when active and passive off-task behaviors were examined separately, but this could have been due to the reduced number of occurrences of each separate behavior type. The presence of a TA also decreased the total amount of pupil off-task behavior. These results indicate that a TA in the classroom can facilitate individual on-task activity and inhibit the overall amount of off-task behavior. The evidence from this study, therefore, is that TAs can have an indirect effect on pupil learning in the sense that their presence benefits interactions between the teacher and pupils.

It is not possible to exactly explain this effect. It might be that the presence of the TA encourages pupils to contribute more, or it could be that when the TA takes over responsibility for some pupils, there are more opportunities for the rest of the class to be involved in interactions with the teacher. In this study, it was not possible to collect detailed information on interactions between TAs and pupils, as the bulk of recorded interactions was between teachers and pupils. Further research, based on close attention to interactions between teachers, TAs, and individual pupils, is needed to better understand the classroom dynamics involved and conceptualize more precisely the pedagogical contribution of TAs in relation to teachers.

General Models of Class Size in Relation to Classroom Processes

In general, therefore, the findings from the present study support the view that small classes lead to more individualized task-related contacts between teacher and pupils—where pupils are less one of the crowd—and a more active role for pupils in interactions with their teachers. It is interesting to note that these findings are at variance with Finn et al.'s (2003) conclusion that class size affects student engagement more than teaching behavior. In the present study, there was no effect of class size on student on- or off-task behavior, and this was also at variance with the earlier English CSPAR study involving pupils aged 4 to 5 years (Blatchford, 2003b). If the age of the student is important in terms of reducing class size effects, as claimed by Finn et al., it seems that it affects student on-task behavior more than teacher–pupil interaction. The present study has shown that effects of class size on teacher–pupil interactions are still evident at 10 to 11 years, and putting these results together with the earlier study of pupils aged 4 to 5 years suggests that this is a stable and robust finding across the primary stage (4–11 years).

This article has reported findings from an observational study of pupil and teacher behavior and cannot address the full coverage of effects contained in general models of class size effects put forward by Anderson (2000), Blatchford (2003a), and Finn et al. (2003). These models include dimensions such as greater knowl-
edge of pupils and in-depth treatment of the curriculum content that cannot be addressed through systematic observation. Connections between class size and other classroom processes have been found in other components of the overall research program and are reported elsewhere (Blatchford, Bassett, et al., 2003).

This article is also limited in providing evidence based on the quantity of teaching behaviors. The study was not designed to measure the quality of teaching in a systematic way, but there were a number of suggestions, from the questionnaires and case studies also carried out in the overall study, concerning ways in which class size could affect the quality and effectiveness of teaching. These were similar to qualitative results from the earlier CSPAR study (Blatchford et al., 2002). In smaller classes, it can be easier for teachers to spot problems and give feedback, set individual targets for pupils, and be more flexible in the use of different styles of teaching.

This study has shown that two other factors need to be added to general models of effects on classroom processes. The first factor is the presence of extra adults, as discussed above. Another factor, as we have shown, is the need to consider ways in which relationships between class size and teacher and pupil behavior can be moderated by the school subject. One direction for future research would be to identify more precisely in which ways in which class size is significant in relation to particular school subjects, as well as factors that might explain any differences found.

Class Size and Pedagogy

As well as considering the context provided by the curriculum when interpreting effects of class size, the overall class size study has suggested that one also needs to consider teachers' pedagogical beliefs. From case studies in classrooms and teacher responses to questionnaires about the effects of class size, we have found that teachers in these English schools have a strong belief in the value of addressing the individual needs of pupils, but this is at variance with curriculum and assessment demands, as we have argued above. This creates considerable tension for some teachers. Teachers' pedagogical views can also conflict with pressures arising from the classroom context, especially large numbers of children. There is an inevitable tension for teachers because of their belief that learning is best served by maximizing individual attention to pupils and also because of the tradition of supporting work to be done individually by pupils. This creates considerable professional concern and is exacerbated when teachers are allocated larger class sizes.

We suggest that one way of addressing this conflict is to do more in teacher training and professional development to consider more strategic linkages between the curriculum, pedagogy, and classroom contexts. More could be done to encourage better use of classroom contexts to help teachers make the most of different class sizes. In the case of large classes, it would be important to develop strategies to maximize the opportunities for individual support. In the case of small classes, it would be important to ensure that the opportunities afforded teachers are taken up. There is evidence that teachers do not always change their teaching approaches when faced with smaller classes (Bertozzi & Randoloph, 1989).

However, we also argue that an alternative solution would be to make more strategic use of a third context for learning to offset the heavy reliance on teacher-led sessions and individual work—that is, more use of group work in the sense of pupils learning together with a deliberate attempt to minimize the teacher's input. With appropriate training for teachers and pupils, this can have benefits for pupils learning and can also help teachers, especially those with large numbers of pupils, in terms of maximizing their time with other pupils and encouraging independence in learning (Blatchford, Kutnick, et al., 2003). It would also help teachers lucky enough to have small classes: As Betts and Shkolnik (1999) found, teachers could make better use of small classes if they did not reduce group instruction.

References


Rowe, K. J. (1985). Factors affecting students’ progress in reading: Key findings from a longitudinal study. Literacy, Teaching and Learning, 1, 57–110.

Received October 13, 2004
Revision received January 20, 2005
Accepted January 27, 2005