Daniel Muijs, Chris Chapman & Paul Armstrong

Teach First: Pedagogy and outcomes. The impact of an alternative certification programme

Abstract

This paper reports on a theory-based evaluation of the Teach First programme, an alternative certification programme based on Teach for America. A mixed methods approach was employed within the theoretical framework of the Dynamic Model of Educational Effectiveness. Findings from classroom observations, interviews and surveys suggest that Teach First teachers in their second year are effective practitioners, using a mainly whole-class interactive teaching approach. Analysis of national student performance datasets using a quasi-experimental design showed that schools partnering Teach First outperformed comparison schools. The study thus provides some support for the effectiveness of Teach First and for the theoretical model.

Keywords

Educational effectiveness; Teacher education; Alternative certification; Teach First

Teach First: Pädagogische Umsetzung und Resultate. Wirkungsweisen eines alternativen Zertifizierungsprogramms

Zusammenfassung

In diesem Artikel werden die Ergebnisse einer theoriebasierten Evaluation von Teach First berichtet, einem auf dem Programm Teach for America basierenden alternativen Zertifizierungsansatz. Im theoretischen Rahmen des Dynamic Model

Prof. Daniel Muijs, PhD (corresponding author), Southampton Education School, University of Southampton, Highfield Campus, Southampton SO17 1BJ, United Kingdom
e-mail: d.muijs@soton.ac.uk

Prof. Chris Chapman, PhD, School of Education, University of Manchester, Oxford Road, M13 9PL Manchester, United Kingdom
e-mail: chris.chapman@Manchester.ac.uk

Paul Armstrong, MA, University of Nottingham, School of Education, Jubilee Campus Wollaton Road, Nottingham NG81BB, United Kingdom
e-mail: paul.armstrong@ioe.ac.uk

of Educational Effectiveness wurde für diese Evaluation ein Mixed-Methods-Ansatz genutzt. Die Ergebnisse aus Unterrichtsbeobachtungen, Interviews und Befragungen deuten darauf hin, dass Teach First-Lehrkräfte im zweiten Jahr effektive Praktiker sind, die vorwiegend interaktiven lehrerzentrierten Unterricht geben. Die Analysen nationaler Schülerleistungsdaten mit quasi-experimentellem Design zeigen, dass Schulen mit Teach First-Partnerschaft Vergleichsschulen in der Leistung übertreffen. Die Ergebnisse der vorliegenden Studie bieten somit einige Fundierung für die Effektivität von Teach First und für das theoretische Modell.

Schlagworte

1. Introduction

1.1 The Teach First programme

The Teach First programme, based on the US programme Teach for America (TfA), was launched in 2002 to encourage high achieving graduates to teach in schools serving low socio-economic status (SES) communities, with the aim of helping to close the achievement gap between high and low SES background students and schools.

Students are assessed by a graduate recruitment company. They then spend six weeks at the Summer Institute, an intensive Summer School, before arriving in schools in September. During the first week they are at a university local to their region. The second week is spent at a school and the third at the secondary school they will be teaching in. They then go to University for weeks four to six, where they learn about classroom management, assessment, resources, special educational needs, social justice, diversity and government policy. All participants make a commitment to be in their training schools for two years (90 % effectively stay on for the two years).

Teach First recruits, trains, places and supports 500–600 teachers per year who are placed in secondary schools in challenging circumstances in three urban conurbations (Greater Manchester, London and West Midlands) in England to teach for at least two-years. For the years this article refers to, 'challenging circumstances' was defined as schools where less than 25 % of young people achieved five GCSEs at Grades A*–C (including English and maths) and/or where at least 30 % of the students were eligible for Free School Meals. The majority of participants teach priority subjects as defined by the Teacher Development Agency (TDA), which include mathematics, English, science, design and technology, information and communication technology, music, religious education and modern foreign languages.

Teach First has been subject to criticism from some commentators (e.g. Smart, Hutchings, Maylor, Mendick, & Menter, 2009; Smyth, 2010), which focuses on what is seen as too limited a training period, potentially leaving teachers underprepared; the two-year term which is seen as exacerbating the problem of staff instability in schools serving disadvantaged communities; the limited recruitment pool, and the private sector sponsorship underlying the project. However, little research has been undertaken on the project. The main evaluation so far is that undertaken by the National Inspectorate, Office for Standards in Education, Children's Services and Skills (Ofsted). Ofsted (2008) reported that the programme made a positive contribution to the schools where the teachers were placed. However, until this study no research existed on the impact of the programme.

Like Teach First, TfA has been subject to various criticisms as well as praise. Criticism of the programme has often focused on pedagogy (Darling-Hammond, 1994). The turnover of teachers has been another source of controversy, though retention appears to be related to the difficulty of the assignments given them, with TfA teachers assigned split grades, multiple subjects, or out-of-field classes more prone to leaving their schools or resigning from teaching (Donaldson & Johnson, 2010).

Given its longer history it is not surprising TfA has been subjected to more extensive research than Teach First. Studies provide a mixed picture with regards to student outcomes. Some studies claim positive effects (Glazerman, Mayer, & Decker, 2006; Decker, Mayer, & Glazerman, 2004), others negative effects (Laczko-Kerr & Berliner, 2002; Darling-Hammond, 1994) and some no effect (Raymond, Fletcher, & Luque, 2001). Some studies have found evidence that TfA teachers' students achieve comparable or better gains in student learning when compared to other similarly experienced teachers in similar schools. Decker, Maye, and Glazerman (2004) conducted a random-assignment study in six Teach for America regions, and found that students of TfA teachers had significantly higher math score gains than those of other novice teachers. Raymond, Fletcher, and Luque (2001) using data from Texas found that TfA teachers outperformed other newly appointed teachers in maths, and performed similarly to more experienced teachers. Xu, Hannaway, and Taylor (2008) used data from North Carolina high school students and found that TfA teachers had a positive effect on student test scores compared to non-TfA teachers. Darling-Hammond, Holtzman, Gatlin, and Heilig (2005) by contrast found that TfA teachers performed similarly to other uncertified teachers while uncertified, and similarly to certified teachers once they themselves had achieved certification. Laczko-Kerr & Berliner (2002) looking at data from five school districts, found that students of under-certified teachers (including TfA teachers) made about 20 % less academic progress than students of regularly certified teachers.

Less is known about the pedagogy of TfA teachers; the above studies do not generally focus on teaching practices of TfA teachers, and do not use classroom observation methods in their studies. Critics have suggested many lack an awareness of learning theory and diversity pedagogy (Darling-Hammond, 1994), while on the

other hand TfA teachers do appear to have a strong sense of self-efficacy in terms of being able to change student outcomes (Smith, 2005). Typically, though, only indirect measures of pedagogy and teacher effectiveness are employed that focus more on the practices used in the training of the teachers than on teachers' own practices. In view of the established relationship between teacher behaviours and student outcomes (Muijs & Reynolds, 2010) this is a major omission, which leaves findings on relationships between TfA teachers and student outcomes as something of a 'black box' of unexplained processes.

1.2 Theory-driven evaluation

In order to explore the impact of Teach First teachers, this paper reports on a theory-driven evaluation of the Teach First programme.

A theory-driven evaluation approach assumes that we can use evaluation methodologies to illuminate theoretical models, linking evaluation of specific programmes to theories of change and action (Creemers, Kyriakides, & Sammons, 2010). All social programmes contain within them an implicit theory of how the world works, and therefore why the change should be beneficial. A key element of theory-driven evaluation is that the evaluation goals and mechanisms are not driven exclusively by stakeholders and evaluation commissioners, but also relate to the theoretical underpinnings that may explain intended outcomes and mechanisms by which they are to be achieved. Work on theories of action has attempted to articulate these underlying mechanisms with a view to improving the evaluation process. A systematic theory of action provides a foundation for programme evaluation, by identifying not only critical program components but also what their logical points of impact will be (Weiss, 1997). A theory of action includes background information; a description of the program components, what the program components intend to achieve, and how they interact; and short- and long term outcomes.

What is, however, still often missing in these theory of action approaches is a clear connection to broader theoretical frameworks that can provide explanations for the success or otherwise of an intervention. Of course, there are many possible theoretical frameworks, the utility of which in part depends on the questions and topics of interest. In the case of Teach First we are primarily interested in impact on students in schools serving low SES communities. This leads us to a need to consider theoretical explanations of the link between school and classroom processes and student outcomes that can explain ways in which employing high level graduates as teachers can be hypothesized to affect outcomes. For this reason we have premised this evaluation on the theoretical framework provided by the Dynamic Model of Educational Effectiveness (Creemers & Kyriakides, 2008).

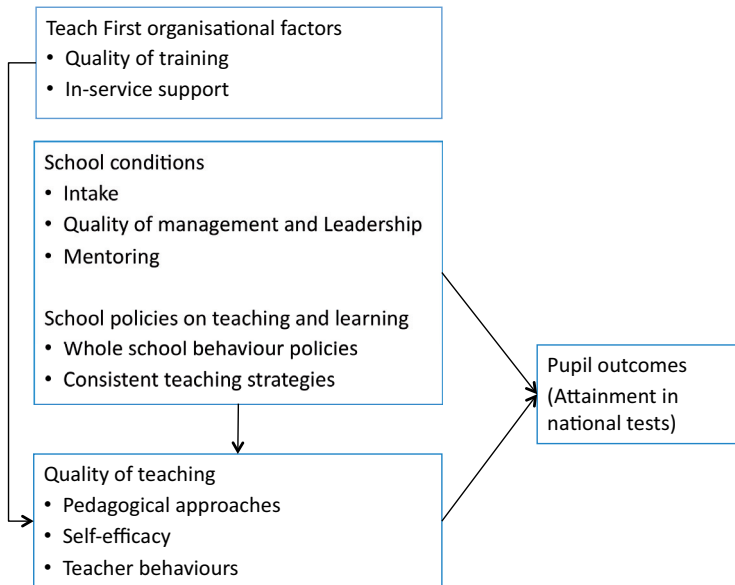## 1.3 The Dynamic Model of Educational Effectiveness

The Dynamic Model of Educational Effectiveness was developed by Creemers and Kyriakides (2008) to build on previous theories in the field of educational effectiveness. A key characteristic of the Dynamic Model is its multilevel nature. Teaching and learning are emphasised, but school-level factors are seen as providing conditions under which the effectiveness of teaching and learning can be maximised by developing and evaluating policies on teaching and creating a positive learning environment in the school. The model also takes into account that the teaching and learning situation is influenced by the wider educational context in which students, teachers, and schools are expected to operate. Factors such as societal values and government policies play an important role in shaping teacher and student expectations.

The model assumes that factors at the school and context level have both direct and indirect effects on student achievement. Therefore, teaching is emphasised and the description of the classroom level refers mainly to the behaviour of the teacher in the classroom and especially to his/her contribution in promoting student learning (Creemers & Kyriakides, 2008).

The dynamic model is based on the assumption that each effectiveness factor can be defined and measured using five dimensions: frequency, focus, stage, quality, and differentiation. Frequency is a quantitative way to measure the functioning of each effectiveness factor. The other four dimensions examine qualitative characteristics of the functioning of the factors. The focus dimension refers to the specificity of the activities associated with the functioning of the factor and the purposes for which an activity takes place. The stage refers to the period and longevity of activities and factors. The quality refers to properties of the specific factor itself, as these are discussed in the literature. Finally, differentiation refers to the extent to which activities associated with a factor are implemented in the same way for all the subjects involved with it.

The DMEE is a useful framework for analysing the impact of Teach First, and provides us with a model proposing a dynamic relationship between school conditions, classroom processes and student outcomes. Of course, not all factors and dimensions of the DMEE are relevant to this study, and we have therefore focussed on those that we hypothesize to be important to evaluating the impact of Teach First. The model we propose is the following:

Figure 1:   The Dynamic Model of Educational Effectiveness applied to this evaluation

```
┌─────────────────────────────────────┐
│ Teach First organisational factors   │
│ • Quality of training                 │
│ • In-service support                  │
└─────────────────────────────────────┘

┌─────────────────────────────────────┐
│ School conditions                     │
│ • Intake                              │
│ • Quality of management and Leadership│            ┌──────────────────┐
│ • Mentoring                           │            │ Pupil outcomes   │
│                                       │            │ (Attainment in   │
│ School policies on teaching and learning│          │ national tests)  │
│ • Whole school behaviour policies     │            └──────────────────┘
│ • Consistent teaching strategies      │
└─────────────────────────────────────┘

┌─────────────────────────────────────┐
│ Quality of teaching                   │
│ • Pedagogical approaches              │
│ • Self-efficacy                       │
│ • Teacher behaviours                  │
└─────────────────────────────────────┘
```

We therefore suggest that the 'black box' between Teach First participation and outcomes can be explained through teaching processes. In particular, we propose that the teachers' self-efficacy (their beliefs in their ability to accomplish a task or goal, in this case improve the performance of students from disadvantaged backgrounds) is related to outcomes, with higher levels of self-efficacy seen as a necessary but not sufficient condition for effective teaching (see, e.g., Bandura, 1998). Teachers' pedagogical approaches and effectiveness in terms of their classroom behaviours have also been found to be related to student outcomes in many studies (see Muijs & Reynolds, 2010) and will therefore be hypothesized to influence outcomes.

Student outcomes, in turn, are expected to be influenced by student background characteristics and overall school characteristics like the effectiveness of the school (see Teddlie & Reynolds, 2000). Quality of teaching is hypothesized to be influenced by school characteristics and external factors such as training and support from Teach First itself.

While this theoretical model doesn't include all possible factors that may influence TF teachers classroom practices or student outcomes (for example, teachers' practices may be influenced by prior educational background or by the makeup of particular classes taught, and student outcomes by factors such as motivation), it attempts to parsimoniously map a number of key theoretical factors within the framework of the Dynamic Model in order to develop understanding of the impact of the Teach First programme in schools. In terms of the dimensions of these

factors, we focus particularly on the quality and frequency dimensions., as these are the elements that are most likely to be impacted by the Teach First programme through its provision and development of non-traditional new teachers who will work in these schools. The model we are testing here therefore provides a partial test of the DMEE in two ways:

- Theoretically, it tests whether specific forms of teacher development may influence the quality and frequency dimensions of educational effectiveness as they relate to the role of classroom teaching.
- Practically, it tests the utility of the DMEE as a tool for programme evaluation, and in particular for the evaluation of interventions in teacher quality.

1.4 Research aims

The key aim of this study was to explore the extent to which Teach First teachers were effective classroom practitioners and could have a positive impact on student learning, and what factors could support them in being effective. On a theoretical level, we were interested in whether this study could provide any additional support to the DMEE.

Research questions therefore were:

- To what extent are Teach First teachers effective classroom practitioners, as perceived by school staff, external observers and colleagues? Do Teach First teachers employ pedagogical approaches which are considered to be effective?
- Is there any evidence that can support a positive impact on learning?
- What factors can help or hinder them becoming effective practitioners?
- Does this evaluation provide support for the DMEE?

2. Methodology

A mixed methods approach was used in this study, as the different variables in our view require different data collection strategies. This design aimed to provide breadth and depth, while ensuring the collection of rigorous and replicable data.

The data collection methods will be outlined below. We collected both quantitative and qualitative data. Quantitative data from surveys, the National Student Database and classroom observation were used to look at the impact of the project on teaching quality and student achievement, while qualitative data from case studies and interviews were used to develop deeper understanding of processes and facilitators and barriers to success.

2.1 Evidence on quality of teaching and school conditions

Data on quality of teaching was collected during case study visits to 16 schools that were taking part in Teach First. Each case study school was visited during the course of the evaluation. A purposive sampling framework was used to select the schools. Three main elements influenced the sampling framework: location, intake diversity and school type. The Teach First project was, at the time of this study, operating in three main areas: London, the urban North-West and the urban West Midlands. Six schools were selected from London, and five each from the other two areas. In terms of intake, while all schools were located in areas of social disadvantage and therefore had a low SES intake, they were diverse with regards to their ethnic intake. We therefore endeavoured to select schools to maximise variance in terms of ethnic intake. Finally, a variety of types of schools exist in urban areas in England, including comprehensive (local authority) schools, Catholic schools, Church of England schools, and academies (independent state schools similar to Charter schools in the US). Again, we aimed to represent all these types in the sample. The final sample was as follows:

Table 1:   The case study sample

Location	Ethnic intake	School Type
London	Predominantly Black	Catholic
London	Varied mix	Academy
London	Varied mix	Comprehensive
London	Black and Asian	Comprehensive
London	Predominantly Asian	Academy
London	Predominantly White	Comprehensive
North-West	Varied mix	Academy
North-West	Predominantly White	Catholic
North-West	Predominantly White	Comprehensive
North-West	Predominantly White	Academy
North-West	Predominantly Asian	Comprehensive
West Midlands	Predominantly Black	Church of England
West Midlands	Varied mix	Comprehensive
West Midlands	Predominantly White	Church of England
West Midlands	Varied mix	Academy
West Midlands	White and Asian	Comprehensive

Schools had between 1 and 9 Teach First teachers working in them, and we studied a total of 47 TF teachers. Two main data collection methods were used in the case study sites: observations and interviews. In each school, the research team conducted classroom observations of Teach First teachers in the case study schools.

A total of 47 lessons were observed. Lessons were videoed to ensure high levels of reliability as it can be hard to rate behaviours reliably on the spot during live observations (Muijs, 2006). Video allows multiple observations of the same lessons to ensure reliable inference. The International Systematic Teacher Observation Framework (ISTOF) classroom observation schedule was used to analyse the classroom observation data. This is an internationally validated rating scale designed by an international team of experts in the area of teacher effectiveness to measure observable teacher behaviours consistent with effective classroom teaching (Teddlie, Creemers, Kyriakides, Muijs, & Yu, 2006). The ISTOF measure is linked to the Dynamic Model of Educational Effectiveness in its attempt to provide a broad and valid measure of the quality of processes that are central to effectiveness in the classroom. One of the key methodological premises of the DMEE is its attention to valid and reliable measurement, and as Creemers and Kyriakides (2008, p. 222) point out, the development of ISTOF followed a concerted process of international validation that serves as a model for the development of other measures linked to the DMEE. The ISTOF Teacher Observation Protocol has 21 indicators spread across seven components of effective teaching. Each indicator is represented by two or three items, resulting in a total number of 45 items (see Table 6). The items are rated on a 5-point Likert scale with values ranging from 'strongly agree' (5) to 'strongly disagree' (1). There is also a 'NA' (not applicable, unable to observe) response option since some of the items may not be relevant or observable in some classroom settings.

Confirmatory factor analysis was undertaken which confirmed the seven component structure. Reliability (Cronbach's Alpha) of the components ranged from .75 to .87. Three raters rated each observation, a satisfactory Cohen's Kappa inter-rater reliability of .83 was achieved.

As well as the classroom observations, we conducted interviews with all second year Teach First teachers in each school ($n = 47$), their head teachers ($n = 16$), their line managers ($n = 31$), and other teachers in the schools ($n = 28$). A semi-structured interview protocol was used. Interviews lasted between 20 and 40 minutes.

Qualitative data collection and data analysis were closely integrated (Miles & Huberman, 1984). This strategy allowed the team to check out hypotheses as they emerged from data analysis and refine data collection strategies as the study progressed. In addition to qualitative analysis, interview data were also analysed using content analytic methods. Content analysis is a summarising, quantitative analysis of messages that relies on the scientific method (including attention to objectivity-inter-subjectivity, a priori design, reliability, validity, generalizability, replicability, and hypothesis testing) and is not limited as to the types of variables that may be measured or the context in which the messages are created or presented (Neuendorf, 2002). A coding scheme was developed and results quantified.

In addition to the case study, a survey was sent out to all second year Teach First teachers each year during the two years of the evaluation. This survey of Teach First teachers also contained a number of questions relating to pedagogy.

In this way we hoped to learn more about teacher behaviours, attitudes and interactions with students. Part of this survey was Bandura's teacher self-efficacy scale (Bandura, 1998). The reliability of this scale in this sample was .75 (Cronbach's Alpha).

2.2 School level impact

A quantitative methodology was used to explore the question of impact of Teach First on student attainment. National student and school level datasets were collected from the Department for Education to allow us to look at performance measures controlled for student background over time. Student Level Annual School Census (PLASC) and National Student Database (NPD) data were requested from, and provided by, the DfE for this purpose. Data were collected for each year from 2001 to 2009.

All Teach First partnering schools were identified through a list provided by Teach First, and their individual school (LAESTAB) number and the year the school first partnered with Teach First were established.

In order to look at the impact of Teach First on performance, we opted for a quasi-experimental design where each Teach First school in the sample was matched to a school as similar as possible on key characteristics prior to joining Teach First. National datasets were used to match schools by

- type of school (e.g., Voluntary Aided, Voluntary Controlled, Academy);
- gender intake (co-educational, single sex boys, single sex girls);
- performance levels (e.g., % achieving KS threshold levels);
- student intake characteristics (% students identified as having Special Educational Needs, percentage students eligible for Free School Meals);
- location; and
- school size (as indicated by student roll).

These data were matched as closely as possible for the three years prior to the school partnering with Teach First using propensity score matching. We then looked at whether Teach First partner schools outperformed those matched schools not partnered with Teach First, with a view towards getting some indication of whether or not Teach First had a positive impact on student achievement.

Overall no significant differences were found between Teach First schools and comparator schools on any of the matching characteristics.

Multilevel statistical models were used to look at the impact of Teach First on performance, measured as number of 5A*–C passes. In the English education system students are assessed nationally at the end of compulsory secondary schooling (after 5 years of secondary education). 5A–C* passes is considered to be a 'good' pass and allows the student easy progression into further education leading to higher education. The percentage of students reaching this threshold is the key performance indicator for schools in the national accountability system, and is

the basis for the national and local 'league tables' of secondary school performance. Levels were school (level 2) and student (level 1). As the data relates to different cohorts in different years analysis of each year was done separately.

In a second phase of the analyses, we explored whether the number of Teach First teachers who had worked in each of the Teach First schools affected the impact of their school. The hypothesis here is that a larger number of Teach First teachers might have a greater impact as a result of a greater impact on school culture, or through the facilitating effect of a critical mass of Teach First teachers as indicated in some of the qualitative data.

Clearly, this method is limited in that it looks at data at the school rather than the teacher level, which makes causality hard to determine and confuses the impact of Teach First with other developments in school. To demonstrate causality three conditions need to be present:

1. The causal variable needs to precede the effect variable in time. This we can demonstrate using our methodology by looking at performance before the schools joined the Teach First programme compared to what happened after they joined.
2. The causal and effect variables need to be correlated with one another. This can be demonstrated by the statistical analyses undertaken.
3. No third variable can be the cause of the relationship demonstrated under condition 2. This is something we cannot demonstrate using this methodology, as the Teach First and comparison schools may differ from one another in ways not captured by the NPD and PLASC data we used. For example, schools opting to partner with Teach First may have more dynamic and/or effective leadership than those that do not.

One key methodological problem we had with this study is that we were not able to access classroom-level data on student attainment. This obviously is a major weakness, as it leaves us unable to clearly connect individual teacher behaviours with student outcomes. The school-level data on student attainment that we do have allow us to compare schools engaged in Teach First with schools not in Teach First. This of course is a very imperfect measure of the impact of the programme, as the data inevitably include many teachers not involved in Teach First. There are also a number of other possible factors that may impact any relationship between TF participation and outcomes, such as differences in the effectiveness of school leadership, school intake characteristics and school effectiveness. These issues cannot be fully resolved. However, we have attempted to strengthen inference in the following ways:

1. A matched sample was constructed that partialled out differences in initial intake characteristics using propensity score matching
2. Data on school characteristics, such as the overall grade Ofsted, the national school inspectorate, had assigned to the school, and the grades for quality of teaching and leadership, were collected from the Ofsted inspection database and included in the analyses

3. Data on the number of TF teachers in each school was collected and included in the analyses.

These measures in themselves of course only partially help us to eliminate other possible causes. The theoretical model does, however, provide a possible explanatory mechanism, but this has to remain tentative while no classroom data are available.

3. Results

3.1 Impact on teaching

3.1.1 Views of Teach First teachers on their impact and effectiveness in teaching

In general, Teach First teachers felt they were able to make a positive contribution to teaching in their schools, following a period of adaptation and induction in their first term. Interviewees felt they were a dynamic presence in lessons, had good subject knowledge and could motivate students. A lot of the positives they perceived in their own teaching were related to dynamism and enthusiasm which they and their colleagues see as motivating for students:

> I'd say I was firm but fair really, I'd like to think that I've got quite a lot of energy, quite creative, I'm really interested in coming up with new lesson ideas, new ways around the topic. (Teach First teacher)

A difficulty some interviewees faced was dealing with the range of student ability. In many of these schools serving disadvantaged communities the range of ability is large, and even set classes can contain very significant variance (Muijs & Reynolds, 2010). In one school which operated mixed abilities in classes the interviewee found this hard at first because "you have to differentiate your teaching" as students had a "huge range of ability". However, as another interviewee stated: "They do prepare you to differentiate in the training we get". Teachers clearly felt that particular strategies were expected of them by their schools: "We're expected to always teach in the way that other teachers do when they get observed". This manifested itself in advice and pressure from mentors and Heads of Department.

Teach First teachers had high expectations of students. Several interviewees, however, mentioned that these needed to be adapted to the reality in which they found themselves, as initial expectations may have been a "bit unrealistic" (Teach First teacher). However, as one interviewee pointed out, the students do step up to the higher expectations and if the teacher then lowers the expectations, they can

meet at a more "realistic" level in the middle. This self-confidence was also demonstrated in the survey. Participants were asked to indicate the extent to which they felt they could influence different aspects in their teaching, using Bandura's teacher self-efficacy scale rated from 1 (no impact) to 9 (very high impact). Results (see Table 2) indicate that respondents tended to see themselves as able to make a difference in all areas, especially in being able to offer alternative explanations and in helping students to value their learning. Respondents were least confident that they could assist families in helping their children to do well, though even for this item the mean score suggests a tendency to see themselves as being able to make at least some difference. Compared to international studies of newly qualified teachers from the US, Canada, Cyprus, Korea, Belgium, the Netherlands, Norway and Hong Kong (mean scores were calculated weighted for sample size) (Woolfolk Hoy & Burke-Spero, 2005; Klassen & Usher, 2010; Muijs & Roe, 1997; Skaalvik & Skaalvik, 2008) Teach First teachers scored higher in most areas, and in particular in motivating students and classroom management factors (2008 survey) and in controlling behaviour and crafting questions (2009 survey). They scored lower on assisting families, possibly due to the highly disadvantaged nature of the schools they were working in.

Table 2:   Teacher self-efficacy scale mean scores

	Mean in TF sample (second year participants 08)	Mean in TF sample (second year participants 09)	Mean of scale among teachers in int'l studies
How much can you do to control disruptive behaviour in the classroom?	6.4	6.9	6.1
How much can you do to motivate students who show low interest in school work?	7.3	6.6	6.4
How much can you do to get students to believe they can do well in schoolwork?	6.9	7.0	6.4
How much can you do to help your students value learning?	7.4	7.0	6.8
To what extent can you craft good questions for your students?	6.8	7.3	6.8
How much can you do to get children to follow classroom rules?	6.7	6.8	6.2
How much can you do to calm a student who is disruptive or noisy?	7.1	6.7	6.4
How well can you establish a classroom management system with each group of students?	7.1	6.8	6.6
How much can you use a variety of assessment strategies?	7.0	7.0	6.7
To what extent can you provide an alternative explanation or example when working with a group of students?	7.6	7.4	7.0
How much can you assist families in helping their children do well in school?	5.4	5.4	6.0
How well can you implement alternative strategies in your classroom?	6.5	6.4	6.5

3.1.2 School views on Teach First teachers

From the interview data it would appear that schools were generally pleased with the teaching skills of Teach First teachers.

Teach First teachers were perceived as having strong subject knowledge, according to interviewees. Initially, they were seen by some interviewees as naïve in the classroom, but they learn quickly: "in the second year you see a tremendous change" (Senior Manager). The first year is seen as challenging for Teach First teachers "but by the end of the second year they have evolved into outstanding teachers" (Head Teacher). The quality of the mentor assigned to the teacher within the school was important in the light of the steep learning curve they were going through, with over half of the interviewed Teach First teachers stating that this was a major determinant of successful integration in their school. Mentoring by university tutors was seen as effective by the majority of respondents.

The teaching practice of Teach First teachers was largely perceived as effective: "Because they're dynamic they teach dynamically, and that always works with students." (Senior Manager).

According to the head teacher survey, Teach First teachers were also seen as being consistent in terms of their quality as classroom practitioners.

Table 3:   Responses to the head teacher survey on variability of quality (percentages)

	Agree strongly		Agree somewhat		Disagree somewhat		Disagree strongly	
	Second year partici- pants 08	Second year partici- pants 09	Second year partici- pants 08	Second year partici- pants 09	Second year partici- pants 08	Second year partici- pants 09	Second year partici- pants 08	Second year partici- pants 09
There is a lot of variance in the quality of Teach First teachers	6.2	8.0	18.8	24.0	50.0	48.0	25.0	20.0

3.1.3 Teacher behaviours

A content analysis was undertaken of the interview data from all interviewees except the Teach First teachers. Keywords from interviews were measured to determine factors that were said to be typical of Teach First teachers. Keywords relating to teacher characteristics were collated, and converted into a percentage of total expressions. The most common are listed in Table 4:

Table 4:   Percentage of keywords relating to Teach First teacher characteristics

Teach First teacher characteristics	%
Listen and learn from other teachers	23.9
Enthusiastic	18.5
Creative	13.7
Not creative	6.4
Hard working	6.4
Resilient	5.9
Energetic	5.8
Adaptive	5.7

85 % of expressions fell in the eight categories above, seven of which were positive, and one of which (not creative) was negative. The latter accounted for just 6.4 % of expressions. The most common expression used was that Teach First teachers listen and learn from other teachers, followed by enthusiasm, and being creative (therefore, Teach First teachers were described as creative more than twice as often as they were described as not creative).

The Teach First teachers appeared to pick up the teaching styles of the schools they worked in: "They very quickly adapt to the styles that are successful in the establishment they're in, and that's been quite clear to me" (Head teacher). The teaching styles of some of the Teach First teachers were described as innovative, creative and confident with a strong presence in the classroom: "Teach First teachers have a lot of creativity and energy which you might not find with teachers who have gone through the traditional route and who maybe stick to tried and tested methods instead of trying new things" (Senior Leader). There was also a willingness to listen to the views of others and act on that advice. One interviewee, for example, believed that there was a misconception amongst some Teach First teachers when they are in training that regular teachers aren't as good as them and are struggling but

> when you get into your placement you realise that it isn't the case at all. In fact, the regular teachers (mostly) are fantastic at their jobs and, not only do you not have any idea what you're doing, but you're being quite arrogant to believe you can come into a school and change things instantaneously. I copy teachers all over the school, especially in terms of behaviour management. But maybe where our strengths lie is in terms of energy because you know that you may only be there for a year or two so if you want to implement a scheme you have to do it now … and because you're only in there for a short time you have a chance to quickly try everything because you have nothing to lose.

Some Teach First teachers felt that their ability to employ innovative teaching methods was lessened by the difficult circumstances in which some of the schools they worked in found themselves. In one school participants experienced what they felt was a lack of encouragement to be innovative in their subjects, mainly because the school was in national challenge (a government initiative to improve the schools in which students are achieving least well in national tests at age 16, with less than 30 % of students achieving 5A*–C grades at GCSE). This led to reluctance on the part of the school to be innovative in case results were affected as the school was under great pressure to improve.

Responses to the participant survey indicated that Teach First teachers perceived their own teaching as having both constructivist and direct instruction elements, though more of the latter than the former (see Table 5). Constructivist teaching emphasises learners constructing their own knowledge through social interaction and realistic tasks, while Direct Instruction focuses on whole-class interactive methods aimed at mastery of small chunks of knowledge before moving on to the next step. Teach First teachers claimed they used constructivist methods like getting students to think about previous lessons, but in general tended towards a structured, teacher-led approach that has been considered to be effective with students in disadvantaged circumstances (e.g., Muijs & Reynolds, 2010; Muijs, Harris, Chapman, Stoll, & Russ, 2004; De Jager, Janssen, & Reezigt, 2005).

Video recordings of classroom teaching were analysed using the ISTOF observation schedule (see Teddlie et al., 2006).

As can be seen in Table A1 (see Appendix), Teach First teachers consistently rated above the midpoint of the scale for the factors observed, indicating overall high levels of teacher behaviours considered effective in the international literature They also rated similarly to an international sample of teachers observed during the construction of the ISTOF instrument (Teddlie et al., 2006), which consisted of experienced as well as novice teachers and was therefore on average more experienced than the TF sample. However, there were clear differences in performance across the different areas. Teach First teachers were particularly strong in creating a positive classroom climate, averaging over 4 on all items. They also rated highly on classroom management, in particular on correcting misbehaviour and minimising disruption, and on instructional skills, with lessons that ran smoothly and followed a logical progression. Where Teach First teachers were somewhat weaker was in promoting active learning and metacognitive skills, rating between 3 and 4 on most items, with the lowest overall rating of 3 being on the item 'The teacher systematically uses material and examples from the students' daily life to illustrate the course content'.

The observations also showed that pedagogies tended to follow a whole-class interactive approach, with fast paced questioning mixing recall and higher order questions, though often more of the former than the latter. There was appropriate use of individual work, and some good use of group work was observed. Contingent praise was used well.

Table 5: Teach First participants views of their own pedagogical approaches (percentages)

	Like me		Not like me		
	Second year participants 08	Second year participants 09	Second year participants 08	Second year participants 09	
When I'm teaching, I make sure I always refer to the content of previous lessons	37.3	35.0	62.7	65.0	When I'm teaching, I get my students to think about previous lessons
It is often necessary to explicitly instruct students so they don't develop misconceptions and don't waste time	59.7	60.8	40.3	39.2	It is always better to let students find out by themselves, so they can construct their own learning
It is better to start with general principles and then give examples	47.7	46.7	52.3	53.3	It is better to start with examples before going onto general principles
I usually get my students to discover what the objectives of the lesson may be through specific challenges and activities	13.4	12.6	86.6	85.8	I usually clearly explain the objectives of lessons myself at the start of the lesson
Students from disadvantaged backgrounds need more opportunities to express themselves in lessons	24.3	32.4	75.8	67.6	Students from disadvantaged backgrounds need more structure in lessons
As a teacher I need to actively instruct students for large parts of the lesson	55.3	56.9	44.7	43.1	As a teacher I am mainly there to facilitate students' group work or individual activities
Clear structures are less important than individual expression for student learning	11.9	13.6	88.1	86.4	Students need clear structures to learn effectively
A high pace is essential, otherwise students will get bored and we won't be able to cover the curriculum	71.2	79.6	28.8	20.4	A slower pace is essential so students can develop a proper understanding of the topic

3.2 Impact on student outcomes

Two-level multilevel models, with students nested within schools, were used to measure the relationship between Teach First status and performance over time. We tested models for each year following schools partnering with Teach First. A

null model was formulated with no predictors. In the next model 'Teach First status' (partnering with Teach First) was added, while in the final model for each year other correlates of achievement were included, such as gender, SEN status, FSM eligibility, age IDACI (Income Deprivation Affecting Children Index) status and ethnicity (majority/minority). Outcome variables were student level GCSE passes at level A*–C.

3.2.1 2003 cohort

For the cohort of schools that had joined Teach First in 2003, only those schools that had participated in the programme for at least 4 of the following 6 years were included in the analyses, making a total of 27 Teach First and 27 comparator schools.

A number of student background variables, notably IDACI status, FSM eligibility and SEN status, were consistently related to outcomes. Teach First status was significantly related to outcomes from 2005 onwards, with students in Teach First schools on average showing higher levels of performance at GCSE. This is suggestive of impact, although other factors, such as prior capacity to change in Teach First partnering as opposed non-partnering schools may of course be a causal factor as well. The correlation of Teach First status with outcomes is quite strong, explaining between 38.9 % (2005) and 46.5 % (2006) of school-level variance in achievement: that is the variance in achievement between students that can be attributed to them attending different schools rather than to individual differences between them.

Table 6: Multilevel models 2003 cohort

	Baseline Model	2004	2005	2006	2007	2008	2009
		A*-C - Coefficient (standard error)	A*-C - Coefficient (standard error)	A*-C - Coefficient (standard error)	A*-C - Coefficient (standard error)	A*-C - Coefficient (standard error)	A*-C - Coefficient (standard error)
Intercept	7.23 (0.32)	8.34 (0.38)	5.60 (1.0)	10.23 (1.8)	10.25 (1.8)	8.69 (0.7)	9.46 (1.7)
Teach First	ns	ns	3.13 (1.4)	3.85 (1.1)	2.50 (1.1)	3.37 (0.9)	3.40 (0.8)
Gender	ns	ns	ns	ns	ns	ns	ns
Age	ns	ns	ns	ns	ns	ns	ns
FSM	-1.01 (0.07)	-0.99 (0.08)	-1.13 (0.09)	-0.89 (0.07)	-1.28 (0.10)	-1.24 (0.12)	-1.09 (0.09)
SEN	-3.78 (0.10)	-4.04 (0.13)	-3.82 (0.12)	-4.13 (0.10)	-3.86 (0.12)	-4.10 (0.14)	-3.97 (0.11)
School Size	ns	ns	ns	ns	ns	ns	ns
Ethnicity	ns	ns	0.52 (0.2)	ns	0.77 (0.3)	0.63 (0.2)	0.48 (0.2)
IDACI status	-4.01 (0.17)	-3.54 (0.24)	-3.70 (0.28)	-3.95 (0.21)	-3.68 (0.18)	-3.71 (0.15)	-3.85 (0.15)
Level 2 pct Variance	15.40	14.24	13.38	11.49	11.51	15.12	15.42
Level 1 pct Variance	84.63	85.77	86.62	88.47	88.52	84.88	84.61

Note. ns = not significant.

3.2.2 2004 cohort

A similar modelling strategy was used for the 2004 cohort, though only 10 Teach First and 10 comparison schools were included in the sample. In this cohort there were again no initial differences between Teach First and non Teach First partnering schools, but from 2007 onwards students in Teach First schools start to outperform students in non partnering schools. Participation in Teach First explains 20.5 % of the variance at the school level in 2007, a percentage that has increased to 35.5 % in 2009. The IDACI code, FSM eligibility, SEN was also a significant predictor of outcomes.

Table 7:   Multilevel models 2004 cohort

	Baseline Model	2005	2006	2007	2008	2009
		A*–C - Coefficient (standard error)	A*–C - Coefficient (standard error)	A*–C - Coefficient (standard error)	A*–C - Coefficient (standard error)	A*–C - Coefficient (standard error)
Intercept	6.8 (1.9)	8.5 (1.6)	8.0 (1.7)	8.0 (1.6)	8.2 (1.6)	9.6 (1.6)
Teach First	ns	ns	ns	5.4 (2.6)	6.6 (2.3)	6.9 (2.3)
Gender	ns	ns	ns	ns	ns	ns
Age	ns	ns	ns	ns	ns	ns
FSM	-1.0 (0.4)	-1.0 (0.4)	-1.1 (0.4)	-1.3 (0.5)	-1.1 (0.4)	-1.2 (0.5)
SEN	-3.5 (0.4)	-3.8 (0.4)	-3.7 (0.4)	-3.7 (0.4)	-3.6 (0.4)	-3.9 (0.4)
IDACI code	-3.6 (0.5)	-3.5 (0.5)	-4.3 (0.5)	-4.5 (0.6)	-4.2 (0.6)	-3.8 (0.6)
Ethnicity	ns	ns	ns	ns	ns	ns
School size	ns	ns	ns	ns	ns	ns
Level 2 percentage Variance	15.4	17.0	16.3	15.7	13.5	12.9
Level 1 percentage Variance	84.6	83.0	83.7	84.3	86.5	87.1

Note. ns = not significant.

3.2.3 2005 cohort

For the 2005 cohort (consisting of 24 TF and 24 comparator schools) there is less evidence of a correlation between achievement and Teach First partnership over time. The only year in which we find a significant Teach First correlation is 2008. IDACI code, FSM eligibility and SEN were consistently significant predictors of GCSE grades.

Table 8:   Multilevel models 2005 cohort

	Baseline Model	2006	2007	2008	2009
		A*–C - Coefficient (standard error)	A*–C - Coefficient (standard error)	A*–C - Coefficient (standard error)	A*–C - Coefficient (standard error)
Intercept	6.9 (0.7)	7.7 (1.2)	7.5 (1.3)	7.8 (1.4)	8.2 (1.2)
Teach First	ns	ns	ns	3.4 (1.2)	ns
Gender	ns	ns	ns	ns	ns
Age	ns	ns	ns	ns	ns
FSM	-0.8 (0.1)	-0.9 (0.1)	-0.8 (0.1)	-1.0 (0.1)	-1.0 (0.1)
SEN	-3.1 (0.1)	-3.3 (0.1)	-3.4 (0.1)	-3.3 (0.1)	-3.2 (0.1)
IDACI code	-3.2 (0.2)	-3.6 (0.2)	-4.0 (0.2)	-3.7 (0.2)	-3.8 (0.2)
Ethnicity	ns	ns	ns	ns	ns
School size	ns	ns	ns	ns	ns
Level 2 percentage Variance	14.2	14.3	15.3	13.5	14.6
Level 1 percentage Variance	85.8	85.7	84.7	86.5	83.4

Note. ns = not significant.

3.2.4 2006 cohort

21 Teach First schools that had been part of the programme for at least two of the following three years, and 21 comparison schools were included in the analyses.

For the 2006 cohort there is some evidence of growing correlations between TF partnering and attainment over time. In this cohort partnering with Teach First is significantly related to outcomes from 2007 onwards, with a strong increase in the strength of the correlation in 2008 and 2009, explaining up to 22 % of between-school variance. The IDACI code, SEN, FSM eligibility and gender were also significant predictors of GCSE grades.

Table 9:   Multilevel models 2006 cohort

	Baseline Model	2007	2008	2009
		A*–C - Coefficient (standard error)	A*–C - Coefficient (standard error)	A*–C - Coefficient (standard error)
Intercept	7.1 (1.6)	9.2 (2.2)	9.1 (1.7)	8.0 (1.6)
Teach First	ns	3.2 (1.4)	5.7 (2.7)	6.5 (2.4)
Gender	-0.6 (0.3)	-0.4 (0.1)	-0.4 (0.1)	ns
Age	ns	ns	ns	ns
FSM	-0.9 (0.2)	-1.1 (0.2)	0.9 (0.2)	0.8 (0.2)
SEN	-2.9 (0.3)	-3.0 (0.4)	-3.2 (0.4)	-3.1 (0.3)
IDACI code	-3.4 (0.4)	-3.2 (0.4)	-3.3 (0.4)	-3.3 (0.5)
Ethnicity	ns	ns	ns	ns
School size	ns	ns	ns	ns
Level 2 percentage Variance	15.9	14.1	12.5	15.7
Level 1 percentage Variance	84.1	85.9	87.5	84.3

Note. ns = not significant.

3.2.5 2007 cohort

For the 2007 cohort, 26 Teach First schools that had been part of the programme for at least two of the following three years, and 26 comparison schools were included in the analyses. Teach First partnering was significantly related to outcomes from 2008 onwards, though the strength of the relationship does not increase over time as it did for the 2006 cohort. Teach First partnering explained approximately 25 % of school-level variance. SEN status, FSM eligibility, IDACI code and gender were also significant.

Table 10:   Multilevel models 2007 cohort

	Baseline Model	2008	2009
		A*–C - Coefficient (standard error)	A*–C - Coefficient (standard error)
Intercept	12.57 (1.6)	10.6 (1.9)	10.4 (1.1)
Teach First	ns	5.1 (2.7)	6.0 (2.2)
Gender	ns	-0.4 (0.1)	-0.3 (0.1)
Age	ns	ns	ns
FSM	-1.2 (0.1)	-1.4 (0.1)	-1.1 (0.1)
SEN	-3.6 (0.1)	-3.6 (0.1)	-3.4 (0.1)
IDACI code	-3.5 (0.2)	-3.8 (0.2)	-3.7 (0.2)
Ethnicity	ns	ns	ns
School size	ns	ns	ns
Level 2 percentage Variance	15.9	14.7	15.5
Level 1 percentage Variance	84.1	85.3	84.5

Note. ns = not significant.

These results are summarised in Table 11 for all cohorts. Years in which there is a significant difference in performance between Teach First partnering schools and non-partnering schools (with Teach First schools showing higher performance levels) are indicated with an X.

Table 11:   When do Teach First schools outperform non Teach First schools (X indicates a statistically significant positive relationship between school partnering with Teach First and pupil attainment)?

Year	Cohort				
	2003	2004	2005	2006	2007
2003					
2004					
2005	X				
2006	X				
2007	X	X		X	X
2008	X	X	X	X	X
2009	X	X		X	X

Overall, there is evidence of a correlation between participation in Teach First and achievement, which appears one to two years following the first year of participation up to 2005, and more quickly in the following cohorts. This relationship is highly significant, typically explaining more than a quarter of the between school variance.

This relationship of course does not imply causality, as mentioned above. Data are for all students in the school, not just those taught by Teach First teachers. Schools may differ in effectiveness, with more dynamic and effective schools possibly taking up the opportunity to take part in Teach First more readily. A variety of intervening factors may have caused the relationship, such as changes in leadership or teacher recruitment. In order to test for intervening variables we conducted an analysis of Ofsted grades for leadership, teaching and overall grades, where again we compared Teach First partner schools and comparison schools. No significant differences were found, and the range of grades within Teach First partner schools did not differ significantly from that within comparison schools or across the sample overall. Nevertheless, other differences not measured by Ofsted may exist. Clearly, however, the pattern exposed here is suggestive and worthy of further study.

3.2.6 Relationship between student outcomes and number of Teach First teachers in the school

In order to further explore the relationship between Teach First participation and student outcomes, we regressed the number of Teach First teachers in partnering schools on the student outcome measure (see GCSE grades). The hypothesis was that a larger number of Teach First teachers might have a greater impact as a result of a greater impact on school culture, or through the facilitating effect of a critical mass of Teach First teachers as indicated in some of the qualitative data. Percentage of students eligible for FSM, percentage of students with SEN, percentage boys and percentage students from ethnic minorities were also entered into the regression models. This was done for every year from 2003 to 2009. Table 12 shows the standardised regression coefficients (Beta). It is important to note that these analyses only refer to those schools that partner with Teach First, and don't include any of the comparator schools. What we have done here is therefore, to calculate the number of Teach First teachers in each Teach First partner school, and correlate this variable with the outcome variable.

Table 12:   Relationship of number of Teach First teachers in the school to pupil outcomes at KS4 (Pearson's r correlation coefficients)

	2003	2004	2005	2006	2007	2008	2009
Percentage of pupils eligible for Free School Meals	-.12	-.10	-.13	-.17	-.11	-.13	-.14
Percentage of pupils with SEN	-.24	-.30	-.24	-.26	-.32	-.30	-.29
Percentage boys	*ns*	-.06	*ns*	*ns*	-.11	-.07	-.05
Percentage pupils from ethnic minorities	*ns*	*ns*	.05	*ns*	*ns*	.06	.05
Number of Teach First teachers	*ns*	.09	*ns*	.13	.14	.11	.13

Note. ns = not significant.

As can be seen in the table, number of Teach First teachers in the school has a significant weak to modest positive relationship with school level outcomes (Cohen, 1988). The strength of the relationship is weaker than that with SEN or FSM percentages (though of a similar order to the latter), but stronger than that with percentage ethnic minorities or boys. It is also clear that the relationship with number of Teach First teachers strengthens in the later years of the project (as these are standardised effect size measures used in the same regression model they can be directly compared).

3.3 Facilitators and barriers to TF effectiveness

Content analysis was conducted on the factors identified by Teach First teachers in interviews as barriers to and facilitators of success. The main terms found were:

Table 13: Key facilitators and barriers

Facilitators	%	Barriers	%
Critical mass of Teach First teachers in school	25.7	Adaptation period	21.8
In-school support	16.5	Lack of in-school support	21.4
Support from Teach First	9.8	Challenging circumstances of school	18.6
Clear and consistent school policies	9.0	Pupils' social background	10.3
Freedom to take initiatives	8.3	Poor pupil behaviour	9.5
Two-year term	7.6		
Good relationships in school	6.6		

Note. % = percentages of the total codings of the content analyses.

The main factors that facilitate success are the presence of a critical mass of Teach First teachers in the school, in-school support and support from Teach First. Clear school policies and freedom to take initiatives are also important. Main barriers to success are the adaptation period in year 1, in that Teach First teachers appear to go through a steep learning curve, especially in semester 1, which limits their effectiveness during that period; lack of in-school support, challenging circumstances in the school, and, linked to that, poor student behaviour.

The first term in school is hard for Teach First teachers, as they initially have some problems adapting to the classroom. As one teacher remarked: "At Christmas, I wanted to leave, but I'm so glad I didn't" (Teach First participant). Another interviewee had concerns about the speed with which Teach First teachers were put in a classroom on their own, questioning whether some candidates may find the situation too much: "I know the early days were very tough, there were tears, there were upsets and a lesser individual may well have found it slightly overwhelming" (Teach First teacher). Participants suggested that schools can work around the programme and help Teach First teachers by organising more "protected time" to ease

them into the first term to a greater extent to allow time for planning and the paperwork of the training course.

As is apparent from the content analysis, the level of support given by the department to which the Teach First teachers were attached played a significant part in their overall experience at the school. In England, most secondary schools are structured around departments, which deliver the teaching of specific subjects, such as Mathematics or English. For example one interviewee was the only teacher in the department because the head of department went on maternity leave. This meant that though she had the opportunity to take on much of the department head role, she also had no departmental support, limiting her effectiveness. Another participant felt constrained because she had joined a department which lacked cohesion and had poor communication so that it did not operate as a department, but consisted of "individual teachers".

As well as formal and informal support to Teach First teachers, clear schoolwide policies and procedures were mentioned as an important facilitating factor by many interviewees. One interviewee commented that: "All the students are treated equally which is good for Teach First teachers because they have a clear set of guidelines to follow and this structure helps them as they don't have much experience" (Senior Leader). A coherent approach and a strong, set order of activities have been found to be important to the achievement of students, especially those vulnerable to school failure such as students eligible for Free School Meals (Creemers & Kyriakides, 2008). Some Teach First teachers felt that not all schools prepared sufficiently for their participation: one interviewee commented that Teach First teachers were eager to take on extra responsibilities and should therefore be encouraged and given scope to do so. She said that many of her friends in other schools had been "battling against the system to get things pushed through, whereas we've been lucky here".

Levels of support from senior management varied between schools. In-school mentoring arrangements and line management were not in all cases strong, and appear inconsistent across schools. This hinders the possible impact Teach First teachers can make by limiting their professional development opportunities in school, especially the opportunity to learn and receive feedback from more experienced teachers. A critical mass of Teach First teachers is also important to their influence in the school. As one Teach First teacher commented:

> We have a big influence because there are quite a lot of us and we're quite young and it's quite a progressive environment anyway so people are open to trying new things. I don't think this is necessarily down to the fact that we're Teach First but more that our school is open minded about change.

Behaviour management was a major training need for many interviewees, especially in the early phases of teaching. According to one interviewee, for example, the main challenge was behaviour and "getting that under control. It was very important from the beginning that I didn't have people walking all over me" (Teach First

teacher). The interviewee had developed a number of behaviour management techniques because "if you can't get the children to behave and listen how can you do anything creative with them in a lesson?".

Some staff in partner schools complained about the two-year term of Teach First teachers, seeing this rapid turnover as somewhat destabilising. However, others felt that this was not untypical for Newly Qualified Teachers more generally. As one head teacher commented "in this city, in any case, a lot of young staff don't stay long, they want to live outside the city, so we are used to that kind of turnover" (Head Teacher). Another interviewee commented that

> while it is true that they are here for only two years, in that two years you get 18 months of absolute quality education, and if they do go out to industry or the professions they are ambassadors with empathy for inner city education. (Senior Manager)

Teach First participants at times saw the attitudes of the school and its management as an impediment to staying on after two years: "if we felt we were valued, and not that we are being exploited, we would probably stay longer" (Teach First Participant).

In order to test whether the barriers and facilitators identified in the content analysis were related to teaching quality a model was tested using a structural equation modelling approach, with total score on the ISTOF scale as outcome variable. Individual ratings for each teacher on the ISTOF scale were summed to construct a total 'effective teaching' score for that teacher, while their responses on the facilitator/barrier questions, as identified through content analysis, were used as the independent variables (operationalized as number of mentions in interviews for each individual).

The model tested assumed the following relationships:

We therefore hypothesized that the extent to which schools are seen as being in challenging circumstances will influence both positive support and lack of support for Teach First teachers, while also directly affecting the effectiveness of teachers as found in a previous study (Muijs & Reynolds, 2003). We hypothesise that if a school is facing challenging circumstances the attention of management will be focussed on dealing directly with these problems, which may reduce attention to providing support for TF teachers. Positive support was hypothesized to be related to effective teaching, the opposite being true of lack of support. Difficulties with adaptation were hypothesized to be negatively related to effective teaching, while positive school policies were hypothesized to be positively related to support. In addition, following the widespread finding in literature of teacher development and school reform that support from management is a key factor in allowing teacher development to occur (Lieberman & Friedrich, 2010), we added support and lack of support as mediating factors in the model. The variables were collected through

Figure 2:   Theoretical model barriers and facilitators
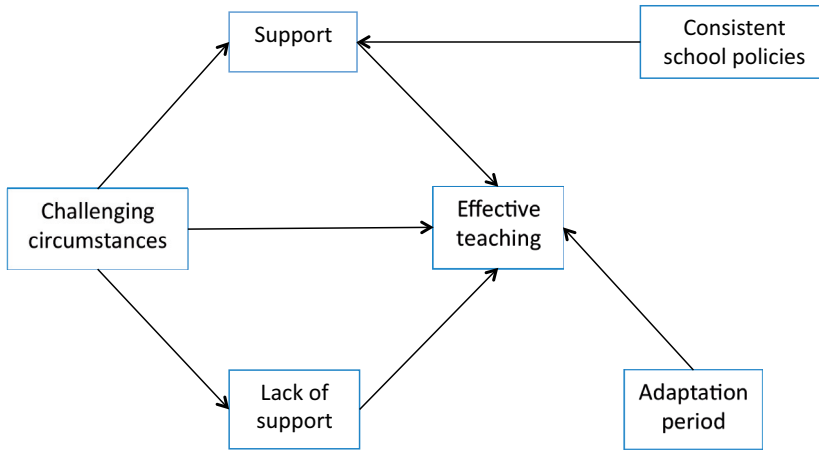
the interview data, where statements regarding positive or negative support were quantified as a percentage of all statements on support.

Model-data fit was tested using a variety of fit indices. The fit indices in Table 14 show acceptable model fit.

Table 14:   Fit Indices

Chi Square	df	RMSEA	CFI	GFI
98.3	36	.46	.97	.98

Some of the predicted relationships were not significant, however. Figure 3 gives all the significant paths and the standardized coefficients. No significant relationship was found between challenging circumstances and positive support, and no direct relationship was found between challenging circumstances and effective teaching.

Figure 3: Significant paths from SEM analysis
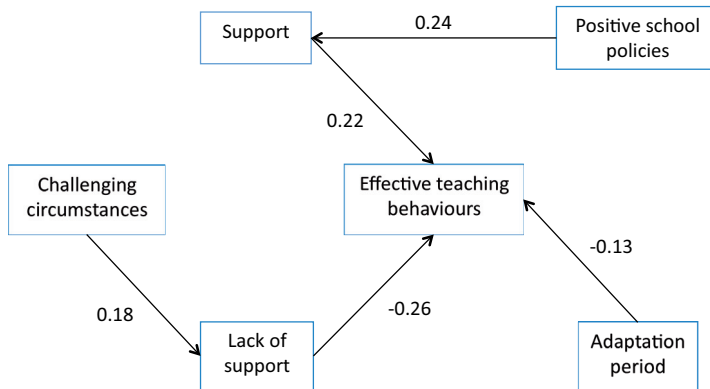
The strongest predictor of effective teaching was lack of support (a negative correlation indicating that lower levels of perceived support are related to less effective teaching). The second most significant predictor was positive support, meaning that where the school supported Teach First teachers strongly they were likely to be more effective; and the third strongest predictor was the year 1 adaptation period. In this case the more an adaptation period was mentioned, the lower the use of effective teaching behaviours. Positive school policies were related to positive support, while challenging circumstances were related to lack of support. School factors thus had a significant indirect relationship with effective teaching by Teach First teachers.

4. Conclusion

What then, can we conclude with regards to our research questions? Our first question related to whether or not TF teachers were good classroom practitioners. The results from this study show high levels of self-efficacy, positive ratings from head teachers, and positive ratings from external observers. Findings therefore are positive in this regard.

Our second research question related to TF teachers' pedagogical practices. Here we find that they tend to follow the most prevalent teaching style in England, whole class interactive teaching based largely on Direct Instruction models (Muijs & Reynolds, 2010). While generally showing behaviours that are supported by prior teacher effectiveness research, there may therefore be a lack of attention to metacognitive instruction and higher order thinking skills.

The results on impact of learning have to remain ambiguous. While we found a positive correlation between schools' participation in Teach First and pupil outcomes, this does not demonstrate causality as various other factors may explain the

difference. Collecting classroom level performance data should therefore be a key aim of future studies.

We identified a number of school and programme level factors that are related to how effective TF teachers can be. These included school-level factors like common policies and support.

Finally, we asked whether the study provides theoretical support to the Dynamic Model of Educational Effectiveness.

In Figure 1 we hypothesized a theoretical model of the possible impact of Teach First on student outcomes, drawing heavily on the Dynamic Model of Educational Effectiveness (Creemers & Kyriakides, 2008). To what extent can we say our data provide support for this model?

Firstly, we hypothesized that Teach First organizational factors would affect teacher behaviours. Here, the evidence is gleaned primarily from the interview data, which suggests that TF teachers feel this factor is important and related to their effectiveness, thus providing some support for this part of the model.

Secondly, we hypothesized that school conditions will influence teacher effectiveness. Again, the data here is taken mainly from the interview data which suggests school conditions like consistency of policies are important. The content and SEM analyses provide support for this part of the model.

Thirdly, we hypothesized that being competent teachers TF participants would contribute to positive student outcomes. Here the evidence comes from a range of sources. There is evidence from the observations and interview data that TF teachers are competent practitioners, with high levels of self-efficacy, though with a generally somewhat traditional teaching methodology. The contention that TF teachers are practitioners using effective teaching behaviours therefore receives support from the data. The link to student outcomes is more tenuous, however. The quasi-experimental design gives us some indication that schools involved in TF may have better student attainment, and that there is a relationship between percentage of TF teachers and outcomes. However, we cannot definitively link this to teacher behaviours found to be effective in the international literature other than to hypothesize that prior research and theory lead us to believe that positive outcomes are likely to be linked to teacher behaviours. This element of the model can therefore only be very tentatively supported.

Fourthly, we hypothesized a relationship between student characteristics and student outcomes. This was supported in the multilevel models.

Overall then, this study provides evidence that Teach First may contribute to school improvement in schools facing challenging circumstances, and thus that alternative approaches to teacher certification, if appropriately designed in terms of recruitment, training and support, can be of benefit as part of a mix of teacher training approaches. It also provides support to the use of the Dynamic Model of Educational Effectiveness as a framework for understanding and evaluating educational reforms, and shows that theory-driven evaluation can enrich our understanding of the phenomena studied. In terms of implications for the development of the model, the findings here point to the importance of the inclusion of teach-

er antecedent characteristics to the model, such as the types and extent of training received, and the educational background of the teachers. It also points to an additional element of school policies to be included in the model, alongside current strands such as consistency in teaching approaches: management support for teaching. In general, the dynamic interplay between school and teacher levels as outlined in the model is supported by this study as being of central importance to Educational Effectiveness.

Of course, further research in this area would be beneficial. In particular collection of classroom-level outcome data and student views, neither of which are included in this study, could lead to more robust findings. We would suggest that future research uses the DMEE to design an intervention study where teachers trained through Teach First are matched to those going through traditional routes, using observation and performance data for students identifiably taught by different teachers are collected. We would also suggest testing students at the start and end of the year to collect change data. Thus, a future study would ideally employ a longitudinal quasi-experimental design, wherein teachers will be followed from their training, through their two years in the project, and beyond into their third year of teaching. So evidence is accumulated year on year about whether the effectiveness of the Teach First teachers develops over time. In the majority of cases, they will be teaching different groups (or even age groups) from year to year, but differences in the initial ability of each group can be taken into account and adjusted for within the statistical analysis. Any increased effectiveness can therefore be measured by comparing the impact of Teach First teachers on pupil attainment in years one, two and three. In order to minimize the effects of other school-level variables, such as school leadership and the impact of other initiatives aimed at reducing educational disadvantage, we would recommend that the comparison groups are located within the same schools using non-Teach First teachers. We would suggest selecting comparison classes from the same year groups as those taught by Teach First teachers. Wherever possible we will use schools that are sufficiently large so that the comparison group will be a parallel class. In this way a robust set of data will be collated.

Acknowledgements

References

Bandura, A. (1998). *Self-efficacy: The exercise of control.* New York, NY: W. H. Freeman and Company.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Creemers, B. P. M., & Kyriakides, L. (2008). *The dynamics of educational effectiveness.* London: Routledge.

Creemers, B. P. M., Kyriakides, L., & Sammons, P. (2010). *Methodological advances in educational effectiveness research.* London: Routledge.

Darling-Hammond, L. (1994). But who will speak for the children? How Teach for America hurts urban school children. *Phi Delta Kappan, 76,* 41–34.

Darling-Hammond, L., Holtzman, D. J., Gatlin, S. J., & Heilig, J. V. (2005). Does teacher preparation matter? Evidence about teacher certification, Teach for America, and teacher effectiveness. *Education Policy Analysis Archives, 13* (42). Retrieved from http://epaa.asu.edu/

De Jager, B., Janssen, N., & Reezigt, G. (2005). The development of metacognition in primary school learning environments. *School Effectiveness and School Improvement, 16* (2), 179–196.

Decker, P. T., Mayer, D. P., & Glazerman, S. (2004). *The effects of Teach For America on students.* Washington, DC: Mathematica Policy Research, Inc.

Donaldson, M. L., & Johnson, S. M. (2010). The price of misassignment: The role of teaching assignments in Teach for America teachers' exit from low-income schools and the teaching profession. *Educational Evaluation and Policy Analysis, 32* (2), 299–323.

Glazerman, S., Mayer, D., & Decker, P. (2006). Alternative routes to teaching: The impact of Teach for America on student achievement and other outcomes. *Journal of Policy Analysis and Management, 25* (1), 75–96.

Klassen, M., & Usher, E. (2010). *Self-efficacy in educational settings.* Bingley: Emerald.

Laczko-Kerr, J., & Berliner, D. (2002). The effectiveness of Teach for American and other under-certified teachers on student achievement: A case of harmful public policy. *Education Policy Analysis Archives, 10* (37). Retrieved from http://epaa.asu.edu/

Lieberman, A., & Friedrich, L. (2010). Teacher leadership: Developing the conditions for learning, support and sustainability. In A. Hargreaves, A. Lieberman, M. Fullan, & D. Hopkins (Eds.), *Second International Handbook of Educational Change* (pp. 647–667). New York, NY: Springer.

Miles M. B., & Huberman A. M. (1984). *Qualitative data analysis: A sourcebook of new methods.* Newbury Park, CA: Sage.

Muijs, D. (2006). Measuring teacher effectiveness. Some methodological reflections. *Educational Research and Evaluation, 12* (1), 53–74.

Muijs, D., Harris, A., Chapman, C., Stoll, L., & Russ, J. (2004). Improving schools in socio-economically disadvantaged areas: An overview of research. *School Effectiveness and School Improvement 15* (2), 149–176.

Muijs, D., & Reynolds, D. (2003). Student background and teacher effects on achievement and attainment in mathematics. *Educational Research and Evaluation, 9* (1), 21–35.

Muijs, D., & Reynolds, D. (2010). *Effective teaching. Evidence and practice* (3rd ed.). London: Sage.

Muijs, D., & Roe, K. (1997). *Literacy in the media age: Results from the third wave of a longitudinal study of children's media use and educational achievement.* Catholic University of Leuven, Department of Communication Science.

Neuendorf, K. A. (2002). *The content analysis guidebook.* Thousand Oaks, CA: Sage Publications.

Ofsted – Office for Standards in Education, Children's Services and Skills. (2008). *Rising to the challenge. A review of the Teach First initial teacher training programme*. London: Ofsted.

Raymond, M., Fletcher, S., & Luque, J. (2001). *Teach for America: An evaluation of teacher differences and student outcomes in Houston, Texas*. Stanford, CA: CREDO, Hoover Institute, Stanford University.

Skaalvik, E., & Skaalvik, S. (2008). Teacher self-efficacy and teacher burnout: A study of relations. *Teaching and Teacher Education, 26* (4), 1059–1069.

Smart, S., Hutchings, M., Maylor, U., Mendick, H., & Menter, I. (2009). Processes of middle-class reproduction in a graduate employment scheme. *Journal of Education and Work, 22* (1), 35–53.

Smith, A. (2005). *Equity within reach: insights from the front lines of America's achievement gap*. New York, NY: Teach for America.

Smyth, J. (2010). Speaking back to educational policy: Why social inclusion will not work for disadvantaged Australian schools. *Critical Studies in Education, 51* (2), 113–128.

Teddlie, C., & Reynolds, D. (Eds.). *The International Handbook of School Effectiveness Research*. London: Falmer Press.

Teddlie, C., Creemers, B., Kyriakides, L., Muijs, D., & Yu, F. (2006) The international system for teacher observation and feedback: Evolution of an international study of teacher effectiveness constructs. *Educational Research and Evaluation, 12* (6), 561–582.

Weiss, C. H. (1997). Theory-based evaluation: Past, present, and future. *New Directions for Evaluation, 76,* 41–55.

Woolfolk Hoy, A., & Burke-Spero, R. (2005). Changes in teacher efficacy during the early years of teaching: A comparison of four measures. *Teaching and Teacher Education, 21* (2), 343–356.

Xu, Z., Hannaway, J., & Taylor, C. (2008). *Making a difference? The effects of Teach for America in high school*. Washington, DC: Urban Institute. Retrieved from http://www.urban.org/url.cfm?ID=411642

Appendix

Table A1: ISTOF mean scores

Category	Indicator	Item	Teach First teachers mean	European studies mean
Assessment and evaluation	The teacher gives explicit, detailed and constructive feedback	The teacher makes explicitly clear why an answer is correct or not	4.3	4.2
		The teacher provides his/her feedback on the answers given by the students	4.4	4.0
	Assessment is aligned with goals and objectives	Assignments given by the teacher are clearly related to what students learned	3.5	3.3
		The teacher explains how assignments are aligned to the learning goals of the lesson	3.2	2.9
Differentiation and inclusion	The teacher creates an environment in which all students are involved	Students communicate frequently with one another on task-oriented issues	3.8	3.7
		Students actively engage in learning	4.2	4.0
	The teacher takes full account of student differences	The teacher makes a distinction in the scope of the assignments for different groups of students	3.5	3.1
		The teacher gives additional opportunities for practice to students who need them	3.4	3.5
Clarity of instruction	The teacher shows good communication skills	The teacher regularly checks for understanding	4.4	4.5
		The teacher communicates in a clear and understandable manner	4.6	4.3
	Clear explanation of purpose	The teacher clearly explained the purposes of the lesson	4.0	3.4
		The teacher asks students to identify the reasons why specific activities take place in the lesson	3.4	2.9
	Lessons are well structured	The teacher presents the lesson with a logical flow that moves from simple to more complex concepts	4.7	4.5
		The teacher implements the lesson smoothly moving from one stage to another with well-managed transition points	4.6	4.6

Category	Indicator	Item	Teach First teachers mean	European studies mean
Instructional skills	The teacher is able to engage students	The teacher provides sufficient wait time and response strategies to involve all types of learners	3.7	4.0
		The teacher gives assignments that stimulate all students to active involvement	4.1	4.0
	The teacher possesses good questioning skills	The teacher poses questions which encourage thinking and elicit feedback	3.5	3.6
		The length of the pause following questions varies according to the difficulty level of questions (e.g., a question calling for application of abstract principles requires a longer pause than a factual question)	3.9	3.6
	The teacher uses various teaching methods and strategies	The teacher uses a variety of instructional strategies during the class period	4.0	3.7
		The teacher uses different strategies for different groups of students	3.2	3.3
Promoting active learning and developing metacognitive skills	The teacher helps pupils develop problem-solving and meta-cognitive strategies	The teacher invites students to use strategies which can help them solve different types of problems	3.6	3.5
		The teacher invites students to explain the different steps of the problem solving strategy which they are using	3.3	3.5
		The teacher explicitly provides instruction in problem-solving strategies	3.1	3.0
	The teacher gives students opportunities to be active learners	The teacher encourages students to ask one another questions and to explain their understanding of topics to one other	3.4	3.7
		The teacher gives students the opportunity to correct their own work	3.4	3.3
		The teacher motivates the students to think about the advantages and disadvantages of certain approaches	4.2	3.8
		The teacher asks the students to reflect on the solutions/answers they gave to problems or questions	3.6	3.5
		The teacher invites the students to give their personal opinion on certain issues	3.2	3.5
	The teacher connects material to students' real world experiences	The teacher systematically uses material and examples from the students' daily life to illustrate the course content	3.0	3.0
		Students are invited to give their own examples	3.8	3.4

Category	Indicator	Item	Teach First teachers mean	European studies mean
Classroom climate	All students are valued	The teacher demonstrates genuine warmth and empathy toward all students in the classroom	4.4	4.5
		The teacher shows respect for the students in both in his/her behaviour and use of language	4.6	4.7
	The teacher initiates active interaction and participation	The teacher creates purposeful activities that engage every student in productive work	4.7	4.6
		The teacher's instruction is interactive (lots of questions and answers)	4.3	4.7
	The teacher interacts with all students	The teacher gives turns to and/or involves those students who do not voluntarily participate in classroom activities	4.5	4.4
		The teacher seeks to engage all students in classroom activities	4.4	4.6
	The teacher communicates high expectations	The teacher praises children for effort towards realizing their potential	4.5	4.4
		The teacher makes clear that all students know that he/she expects their best efforts in the classroom	4.6	4.4
Classroom management	Learning time is maximized	Teacher starts lesson on time	5.0	4.6
		Teacher makes sure that students are involved in learning activities until the end of the lesson	4.7	4.3
		Actions are taken to minimize disruption	4.9	4.5
	Clear rules are evident	There is clarity about when and how students can get help	4.4	4.4
		There is clarity about what options are available when the students finish their assignments	3.3	3.8
	Misbehaviours and disruptions are effectively dealt with	The teacher corrects misbehaviour with measures that fit the seriousness of the misconduct (e.g., she does not overreact)	4.7	4.6
		The teacher deals with misbehaviour and disruptions by referring to the established rules of the classroom	4.1	3.8