Evaluation of the implementation and impact of an integrated prevention model on the academic progress of students with disabilities.

DOI:
10.1016/j.ridd.2014.10.029

Document Version
Final published version

Link to publication record in Manchester Research Explorer

Citation for published version (APA):

Published in:
Research in developmental disabilities

Citing this paper
Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights
Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy
If you believe that this document breaches copyright please refer to the University of Manchester’s Takedown Procedures [http://man.ac.uk/04Y6Bo] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.
Evaluation of the implementation and impact of an integrated prevention model on the academic progress of students with disabilities

Alexandra Barlow *, Neil Humphrey, Ann Lendrum, Michael Wigelsworth, Garry Squires

Manchester Institute of Education, University of Manchester, UK

A R T I C L E   I N F O

Article history:
Received 31 July 2014
Received in revised form 14 October 2014
Accepted 15 October 2014
Available online 12 November 2014

Keywords:
Inclusion
Academic achievement
Achievement for All
Evaluation
Implementation

A B S T R A C T

In this paper we report on the implementation and impact of an integrated prevention model (Achievement for All – AfA) to improve the educational experiences and outcomes of students with disabilities. It comprises three inter-related strands: assessment, tracking and intervention; structured conversations with parents; and, developing provision for wider outcomes. Participants were 12,038 students with disabilities from 431 mainstream primary and secondary schools across 10 Local Authorities in England involved in the two-year AfA pilot. Pre- and post-test data on academic attainment in English and Maths were compared with national data on academic progress for students with and without disabilities over an equivalent period of time. School-level contextual and implementation data and student-level socio-demographic and psychosocial data were also collected. Four hypotheses were tested regarding the impact of AfA on academic attainment in English (H1) and Maths (H2); the influence of aspects of the implementation context and processes (H3); and individual differences between students (H4). Our findings are discussed in relation to the identification and validation of critical intervention components and standards for assessing the practical significance of attempts to improve outcomes for students.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Students with disabilities represent a significant proportion of the school-aged population, and are widely considered to be the most vulnerable group of learners in any education system. Prevalence estimates vary by country – for example, 21% in England (Department for Education, 2010), 13.2% in the United States (Institute of Education Sciences, 2011), and 7.6% in Australia (Australian Institute of Health & Welfare, 2004). This variation is accounted for by differential definitions of disability, and the social, political, historical and legal influences that affect schooling systems across the world (Robson, 2005; Squires, 2012). These influences also have a bearing on how students’ needs are classified. The overwhelming majority of students with disabilities attend mainstream (‘regular’) schools in developed nations (for example, 91.4% in England – Department for Education, 2010; 95% in the United States – Institute of Education Sciences, 2011).

* Corresponding author at: University of Manchester, Manchester Institute of Education, School of Environment, Education and Development, Oxford Road, Manchester M13 9PL, UK. Tel.: +44 161 275 3504.
E-mail address: alexandra.barlow@manchester.ac.uk (A. Barlow).

http://dx.doi.org/10.1016/j.ridd.2014.10.029
0891-4222/© 2014 Elsevier Ltd. All rights reserved.
Promoting positive outcomes for students with disabilities has been perhaps the most significant challenge facing schools in recent years. Despite additional support, the available evidence suggests that they perform less well academically than those without disabilities at each stage of their education in core academic subjects such as mathematics and science (Department for Education, 2010; Zhang, Katsiyanis, & Kortering, 2007). Research also points to poorer outcomes in psychosocial domains, such as: being over-represented as victims (and in some cases, as perpetrators) of bullying (Monchý, Pijl, & Zandberg, 2004; Van Cleave & Davis, 2006); and having lower peer acceptance and fewer friends/friendship groups than their non-disabled peers (Frostad & Pijl, 2007; Pijl, Frostad, & Flem, 2008). These outcomes may be heavily mediated by parental engagement. Unfortunately, this is another area of concern for students with disabilities (Lamb, 2009; Lendrum, Barlow, & Humphrey, 2013).

2. School-based interventions for students with disabilities and the potential of integrated models of prevention

Examination of the literature on school-based interventions for students with disabilities highlights a range of issues that are pertinent to the current study. While a range of strategies and approaches have been shown to be effective, for example, Kavale’s (2007) meta-analysis, interventions are often narrowly-focused, typically targeting a single group (e.g., students with autism) and/or outcome domain (e.g., social skills). This is known as the ‘program for every problem’ phenomenon. While tailoring to individual needs can of course be beneficial, it can also lead to a fragmented approach to implementation that is neither cost-efficient nor sustainable. This means that some very promising approaches are not enacted consistently over time. What is perhaps needed is an integrated model, in which independent strategies or programs at both universal and targeted/indicated levels are fused into one coherent framework for intervention in a coordinated manner (Domitrovich et al., 2010).

An additional problem is the ‘research to practice gap’. Schools rarely use approaches that have a strong evidence base (Cook & Cook, 2011). Even when they do, they typically modify an intervention to improve fit with their own needs and practices (Harn, Parisi, & Stoolmiller, 2013). It is clear that some modifications risk affecting the integrity of the intervention and the achievement of expected outcomes (Durlak & DuPre, 2008; Lendrum & Humphrey, 2012). However, despite an increased emphasis on implementation science in recent years, there is still a clear lag in terms of identifying and operationalizing what actually makes interventions work – with very few studies establishing what are typically referred to as ‘core components’ or ‘active ingredients’ (Blase & Fixsen, 2013). Hence, an intended contribution of the current study was to document the relationship between theorized ‘active ingredients’ and outcomes (see Hypotheses 3a and 3b below). Additionally, we sought to explore the influence on outcomes of factors affecting implementation, such as school leadership support (Gottfredson & Gottfredson, 2002).

Finally, the interpretation of outcome data can be problematic – interventions are often quoted as producing effect sizes that are arguably arbitrary when not viewed in context. Setting appropriate benchmarks for the practical significance of intervention effects in making a noticeable difference in the classroom is becoming an increasingly important consideration (Hill, Bloom, Black, & Lipsey, 2008). Furthermore, the importance of individual differences should not be underestimated. Although traditional approaches to analyzing the impact of interventions (e.g. intention to treat – Gupta, 2011) may imply uniform effects across participants, this is rarely the case. Indeed, developmental theory and research suggests considerable variability in the degree to which individuals are affected by factors in their environment (including interventions) (Belsky & Pluess, 2009). Hence, the influence of student characteristics including (but not limited to) disability type, special educational provision, gender and age should be assessed in order to account for possible ‘differential gains’ (Humphrey, 2013) among sub-groups when examining program impacts (see Hypothesis 4). Such analyses also serve to increase our understanding of the extent to which interventions address inequities – as Petticrew et al. (2012) state, “policymakers and practitioners need evidence on the effects of interventions in subpopulations” (p.97).

The intervention examined in this paper aligns with the ‘integrated prevention model’ outlined by Domitrovich et al. (2010). This theoretical model reflects the complexities of the etiological factors underpinning outcomes for students with disabilities, positing that a range of individual and environmental factors can place children at risk of negative outcomes, and that these outcomes are inter-related. Domitrovich et al. (2010) propose that approaches to intervention that focus narrowly on a single risk factor or outcome domain are less likely to be successful than those that target multiple factors. They argue that effective school-based prevention should combine universal, school-wide approaches with targeted/indicated intervention for specific groups of students. Finally, they suggest that an integrated model, in which independent strategies or programs are fused into one coherent framework for intervention in a coordinated manner, will result in a synergistic effect. The rationale for such a model is provided in terms of (a) comprehensiveness; (b) maximized intervention exposure; (c) the additive or multiplicative effects caused by the interaction of different strategies; (d) reduced ‘initiative overload’ and improved sustainability; and (e) improved potential for high quality implementation.

3. The Achievement for All (AfA) intervention framework

AfA was developed by the English government’s Department for Education in response to on-going concerns about the school experiences and outcomes of students with disabilities, and was conceptualized as a means to support schools to provide better opportunities for them to fulfill their potential. The pilot version was implemented from 2009 to 2011 in over 450 schools across 10 Local Authorities (akin to school districts) in England, supported by a £31 (c. $46.5) million grant.
In the interests of clarity, we adapt the Template for Intervention Description and Replication (TIDieR – Hoffmann et al., 2014) to describe the AfA model:

1. **Brief name**

   Achievement for All (AfA)

2. **Why: Rationale, theory and/or goal of essential elements of the intervention**

   The aim of AfA is to improve the academic achievement, parental engagement and confidence, and wider outcomes (e.g. behavior, positive relationships) of students with disabilities. The pilot program comprised three inter-related strands. Strand 1, ‘Assessment, tracking and intervention’, focused on academics and included the use of the Assessing Pupil Progress1 approach to track students’ progress, the setting of curriculum targets and implementation of appropriate interventions to support children and young people with disabilities to make progress in their learning (Department for Children Schools and Families, 2009a, 2009b; Office for Standards in Education, 2011). Strand 2, ‘Structured conversations with parents’, focused on the use of a clear framework for developing an open, on-going dialogue with parents about their child’s learning and life in school. Training was provided for schools, which emphasized the building of parental engagement and confidence via a four-stage model (explore, focus, plan, review) in three SCPs per year. Finally, Strand 3, ‘Developing provision for wider outcomes’, involved schools developing or enhancing school-wide strategies and key actions to support students with disabilities to make progress in their attendance, improve behavior, reduce exposure to bullying, develop positive relationships and increase wider participation in school life. Strategic support for implementation was provided by professionals working at local (AfA leads and advisory teachers) and regional (National Strategies coordinators) levels.

   In terms of theoretical base, AfA aligns with Domitrovich et al.’s (2010) integrated prevention model in several respects:

   - Reflecting the need to address multiple inter-related risk factors and outcomes for students with disabilities;
   - Presenting a synergy between school-wide strategies (for example, the Developing Provision for Wider Outcomes strand) and targeted provision (the Assessment, Tracking and Intervention and Structured Conversations with Parents strands);
   - Providing a framework for fusing independent strategies and programs in a coordinated manner.

   It reflects aspects of the school-based prevention evidence base that are known to promote high quality implementation and ultimately, sustainability. So, for example, it provides a balance between prescriptiveness (e.g. the Structured Conversations with Parents strand) and flexibility (e.g. the Developing Provision for Wider Outcomes strand) that can secure consistent delivery that is responsive to local context and need (Castro, Barrera, & Martinez, 2004). AfA also involved training for school staff in key aspects of program delivery (e.g. conducting structured conversation with parents) in addition to ongoing ‘support and challenge’ visits from the aforementioned external professionals, akin to technical support and assistance. Both training and technical assistance have been associated with improved implementation in the prevention literature (Durlak & DuPre, 2008; Greenberg, Domitrovich, Graczyk, & Zins, 2005; Joyce & Showers, 2002).

   In the interests of clarity, the program logic model is presented as Fig. 1.

   Inputs in the three strands are designed to systematically address the multiple risk factors to which students with disabilities may be exposed in varying degrees (see above). High quality implementation of these strands is facilitated by the fact that they are deeply inter-related, by the balance struck between universal and targeted provision, and the focus on both flexibility and prescription. Where appropriate, as in the Structured Conversations with parents strand, implementation with high levels of adherence (to program guidance) and dosage (number of conversations completed) is considered a critical component of delivery. These processes are hypothesized to trigger multiple mechanisms of change, including improvements in school-home relations, a comprehensive approach to intervention tailored to individual needs, and an interactive, synergistic effect brought about by multiple complementary approaches to intervention. In turn, these mechanisms bring about improved outcomes for learners with disabilities in both academic and psychosocial domains. Consistent with the underlying theoretical model, temporal relations between these outcomes are expected.

   The AfA logic model recognizes that interventions do not occur in a vacuum and as such two sets of factors affecting implementation are proposed, relating to the school environment and support structures and mechanisms within it respectively. In relation to the former, observed (e.g. proportion of students with disabilities) and latent (e.g. climate, ethos) contextual features of the school are assumed to play a role in influencing the implementation variability of the three program strands, the processes and mechanisms of change, and ultimately outcomes for students. For example, a higher proportion of students with disabilities in a given school may influence the responsive capacity of staff, particularly for

---

1 Assessing Pupil Progress is a structured approach to in-school assessment that, “enables teachers to make judgments about their pupils’ attainment, keyed into national standards, develops and refines teachers’ understanding of progression in their subject…[and] enables teachers to track pupils’ progress over time” (DCSF, 2009, p.2). It involves a three step approach to assessment, namely (i) consider evidence (e.g. from students’ written work and oral contributions), (ii) review evidence (e.g. mapping aspects of the range of students’ work to assessment criteria), and (iii) making a judgment (e.g. about different aspects or assessment focuses within a subject, or an overall subject-level judgment). The approach was found to increase, “teachers’ accountability and the accuracy and consistency of assessment practice” (OFSTED, 2011, p.2).
targeted elements of the program. With reference to the latter, factors such as the involvement of school leadership and external support are considered to be crucial supports for high quality implementation. These proposals are consistent with the implementation science literature (e.g. Durlak & DuPre, 2008; Forman, Olin, Hoagwood, & Crowe, 2009; Greenberg et al., 2005) but are rarely included in program logic models (or, indeed, tested empirically).

3. **Who: Recipients of the intervention**

Students with disabilities are the primary intervention recipients. In the English education system, disability categories are as follows: specific learning difficulties; moderate learning difficulties; severe learning difficulties; profound and multiple learning difficulties; autism spectrum disorder; speech, language and communication difficulties; behavioral, emotional and social difficulties; physical disability; hearing impairment; visual impairment; and, multi-sensory impairment. Provision for students with disabilities escalates as follows: School Action, in which teachers make reasonable adjustments in order to meet a student’s needs as part of their differentiated practice; School Action Plus, characterized by the involvement of external agencies (e.g. educational psychologists), who provide expert support and advice; and, Statement of Special Educational Needs, a legal document outlining support needs that is prepared following a comprehensive, multi-agency assessment process and typically secures additional financial resources to enable these to be met (Department for Education, 2014).

4. **What (materials): Physical or informational materials used in the intervention.**

All participating schools and their Local Authorities received informational materials relating to the AfA program. These included:

- Program prospectus for Local Authorities (Department for Children Schools and Families, 2009b);
- Program guidance for schools (Department for Children Schools and Families, 2009a);
- Strand-specific program guidance (e.g. structured conversations manual – National Strategies, 2009).

5. **What (procedures): The procedures, activities and/or processes used in the intervention**

The delivery model for the intervention was as follows:

---

**Fig. 1. Achievement for All logic model.**
Strand 1: Assessment, Tracking and Intervention

- Use of Assessing Pupil Progress and Progression Guidance to track students’ progress;
- Setting of curriculum targets and implementation of appropriate interventions to support children and young people to work toward those targets. These interventions typically comprised a blend of 1:1 tuition, adjustment of curriculum planning and teaching strategies to address gaps in learning, and use of specialist resources or guided group work to construct inclusive learning sequences;
- Sharing of academic targets and data with parents, through structured conversations (see below) and other media (e.g. emails, text messages);
- Adaptation of interventions based upon frequent target/progress reviews;
- Assessment and tracking at class and school level, supported by meetings to share levels of progress within schools among teachers and management;
- Use of Information Management Systems (such as ‘Assessment Manager’) to support the above.

Strand 2: Structured Conversations with Parents

- A target of 3 structured conversations per school year involving the Key Teacher\(^2\) and parents/carers of every student with disabilities in the study cohort. Where appropriate and feasible, students’ views were also gathered;
- Each structured conversation scheduled for approximately 40–60 min;
- Drawing up of a contract or agreement with parents/carers about the purpose and outcomes of the conversations;
- Utilization of a four stage model in each conversation, as follows:
  - Explore – active listening, paraphrasing, communicating, understanding;
  - Focus – identifying priorities, clarifying key issues;
  - Plan – agreeing targets, seeking outcomes, developing a plan;
  - Review – summarizing, clarifying next steps, arranging further communication.
- The area(s) of focus in each structured conversation to be mutually determined by the Key Teacher and parents/carers, but always directly pertaining to the students’ progress in school and any issues influencing this.

Strand 3: Developing Provision for Wider Outcomes

- Schools required to choose two of five wider outcomes (based on local need) around which to develop provision, introducing or extending school-wide and targeted interventions designed to support students to attend school more frequently, improve behavior, reduce exposure to bullying, develop positive relationships with teachers and peers and/or increase wider participation in school life (e.g. after-school clubs);
- Examples of attendance initiatives (implemented by 41% of schools): reward systems, increased communication with parents (e.g. through structured conversations), using a school attendance officer, and making the curriculum more relevant/accessible to students;
- Examples of behavior initiatives (implemented by 37% of schools): behavioral expectations systems, staff training, reward systems and individual/small group interventions;
- Examples of bullying initiatives (implemented by 5% of schools): peer-based techniques (e.g. mentoring), restorative justice approaches, and use of social and emotional learning materials;
- Examples of positive relationships initiatives (implemented by 63% of schools): Circle Time, peer-based approaches (e.g. buddy benches), and social-emotional learning activities.
- Examples of wider participation initiatives (implemented by 54% of schools): provision of ICT facilities for students and their parents, private and voluntary sector projects, and developing extended services projects specifically for targeted students.

6. Who (provider): Intervention implementers

Staff directly involved in the delivery of AfA included Key Teachers (see above) of students with disabilities, support staff (e.g. teaching assistants), each school’s special education coordinator, and the head teacher. The nature of AfA meant that other members of school staff were also involved, albeit indirectly (for example, other members of teaching staff would support the delivery of the school-wide aspects of the Developing Provision for Wider Outcomes strand). Finally, some aspects of the Developing Provision for Wider Outcomes strand necessitated input from external professionals (for example, a school attendance officer).

Two staff (typically the head-teacher and special education coordinator) from each participating school attended initial ‘launch conferences’ that were held in each of the 10 Local Authorities. These conferences served as a means to introduce AfA and provide some initial training around key principles of the program. These were followed by more specific training for

\(^2\) The ‘Key Teacher’ was defined as the member of teaching staff who knew the individual student best; typically this was their class teacher in primary schools and their form tutor in secondary schools.
staff in specific aspects of implementation (e.g. structured conversation and Assessing Pupil Progress training). Further training was then cascaded in each school to Key Teachers of students in the study cohort.

7. How: Mode of delivery

AfA was primarily implemented using a face-to-face delivery modality, although certain elements utilized other methods where appropriate (for example, strategies for improving assessment, tracking and intervention included the use of email, the Internet, and/or text messages to communicate information about student progress). The Assessment, Tracking and Intervention strand was provided individually or in small groups. The Structured Conversations with Parents strand was provided individually. Finally, the Developing Provision for Wider Outcomes Strand was delivered using a combination of school-wide, small group and/or individual methods.

8. Where: The location of the intervention

AfA was delivered on-site in participating schools.

9. When and how much: Duration and dosage of the intervention

AfA was implemented over the course of two school years. The nature of the intervention (e.g. not an ‘annualized’ series of sessions that could be easily quantified) means that precise dosage is difficult to specify; indeed, participating schools were expected to be actively implementing all three strands throughout the evaluation period. Exceptions to this included explicit guidance to conduct three structured conversations with parents per student per year (Strand 2), and a target of at least termly (every c.13 weeks) review of academic targets and the interventions in place to enable them to be met (Strand 1).

10. Tailoring: Adaptation of the intervention

The AfA intervention model was designed to provide a balance between flexibility and prescription. The Assessment, Tracking and Intervention and Developing Provision for Wider Outcomes strands thus deliberately and explicitly offered scope for adaptation. In relation to the former, although the use of Assessing Pupil Progress and Progression Guidance to inform student tracking was mandatory, the nature of interventions to improve their academic progress (see above) necessarily varied within and between schools and over time, particularly given that an ‘adaptive intervention’ approach (Fabiano, Chafouleas, Weist, Sumi, & Humphrey, 2014) was being encouraged (e.g. if a termly review indicated that a given strategy was proving ineffective, staff were expected to implement an alternative intervention). In relation to the latter, schools were encouraged to focus on the two wider outcomes that best reflected local need. The AfA guidance included content on effective forms of intervention for each wider outcome. This was supplemented by support and advice from the Local Authority AfA lead and advisory teachers.

11. Implementation: Planned and actual

Each school worked in collaboration with their Local Authority AfA lead and advisory teachers to develop, implement and monitor an action plan for delivery of the program based upon the principles outlined above. Additionally, the research team assessed the implementation of AfA as part of the independent evaluation (see below). Data generated through this process confirmed that the majority of participating schools implemented AfA as planned. For example, all but one school reported using Assessing Pupil Progress, with 69% reviewing academic targets on an at least termly basis (Strand 1). Similarly, the average reported fidelity rating for the implementation of structured conversations with parents was 30 (on a scale of 0–36, with higher scores indicating greater adherence to program guidance).

4. The current study

In 2009 the authors were commissioned by the Department for Education to provide an independent evaluation of implementation and impact of the AfA pilot. The evaluation was driven by two objectives. First, to determine the impact of AfA on a variety of outcomes for students with disabilities. Second, to identify the processes and practices that were most effective in improving these outcomes, thereby allowing for the identification and validation of critical intervention components.

This paper focuses on the implementation and impact of AfA on academic outcomes (specifically, attainment in English and Maths) for students with disabilities. Articles on the Structured Conversations with Parents and Developing Provision for Wider Outcomes strands have already been published. Lendrum et al. (2013) found increases in parental engagement and confidence (albeit not to a statistically significant degree) as a result of participation in the AfA program. While the very limited parental response rate (N = 294) and profile of respondents (those parents who completed our measures consistently were, by definition, more actively engaged, essentially leading to a baseline ‘ceiling effect’) made it difficult to draw firm conclusions, qualitative data gathered from schools and parents indicated that this strand was perceived as extremely
successful in building positive school-home relationships. Parental engagement and confidence outcomes were found to vary by disability type, ethnicity, age and special educational provision at the individual level.

Humphrey, Lendrum, Barlow, Wigelsworth, and Squires’ (2013) analysis indicated that participation in AfA was also associated with significant increases in positive relationships and reductions in bullying and behavior problems among students with disabilities. These outcomes varied as a function of a range of aspects of the implementation environment (e.g., the proportion of students from low socio-economic backgrounds), implementation processes and practices (e.g., dosage and/or fidelity to the Structured Conversations with Parents model), and student characteristics (e.g., sex). Three reports to the funder are also available (Humphrey & Squires, 2011a,b; Humphrey & Squires, 2010, 2011b).

The above studies indicate that the AfA pilot was largely successful in improving a range of salient outcomes for students with disabilities. However, consistent with the program logic model and theory and research outlined in Section 2, the magnitude of impact was influenced by variability in contextual features of schools, their implementation practices, and differences in student characteristics.

4.1. Setting benchmarks for practical significance

A complicating factor in examining the impact of an intervention is that statistically significant treatment effects may not always be practically significant and make a noticeable difference in the classroom. In light of this, in the current study we draw upon the work of Hill et al. (2008). They suggest 3 benchmarks – normative expectations for change (that is, how does the effect of an intervention compare to an equivalent period of growth for a given target population of learners?), policy-relevant performance gaps (that is, how does the effect of an intervention compare with existing differences among particular groups of learners, such as those with and without disabilities?), and observed effect sizes for similar interventions (that is, how do the effects of an intervention compare to those from previous studies?). We predicted that exposure to the AfA intervention would lead to significantly greater progress in English (Hypothesis 1a) and Maths (Hypothesis 1b) for students with disabilities in pilot schools than students with disabilities nationally over the equivalent period of time (Benchmark 1). However, a more stringent benchmark – and one that would allow us to ascertain whether AfA was successful in ‘narrowing the attainment gap’ – was the prediction that students with disabilities in intervention schools would also make more progress in English (Hypothesis 2a) and Maths (Hypothesis 2b) than students without disabilities nationally (Benchmark 2). For Benchmark 3, a problematic issue is that most meta-analyses reporting on effect sizes of interventions and influences on academic attainment (e.g. Hattie, 2009; Hill et al., 2008) focus on the general population of students as opposed to those with disabilities. However, Swanson and Hoskyn’s (1998) meta-analysis does provide information pertaining specifically to effect sizes of interventions for students with disabilities. Although the interventions included in their analysis were more intensive in nature than the AfA pilot, the average effect size of \( d = 0.45 \) on general academic achievement nonetheless provides a useful threshold for comparison.

In light of the body of evidence presented earlier in this article – and consistent with the AfA logic model (Fig. 1) and our findings reported elsewhere (e.g., Humphrey, Wigelsworth, Barlow, & Squires, 2013) – we also predicted that variability in key aspects of the implementation environment (Hypothesis 3a) and implementation processes and practices (Hypothesis 3b) would predict variability in intervention outcomes. Finally, in recognition of the emerging body of evidence pertaining to differential responsiveness to intervention, we predicted that multiple individual differences between students would account for significance proportions of variance in our study outcomes (Hypothesis 4).

5. Method

5.1. Design

The current study comprised two main components:

- A quasi-experimental, pre-test-post-test design to determine the impact of AfA on academic progress when compared with children with and without disabilities nationally (Hypothesis 1a, 1b, 2a, 2b);
- A longitudinal, multi-level, natural variation design to determine the relative influence of different aspects of the implementation environment (Hypothesis 3a), implementation processes and practices (Hypothesis 3b), and individual differences between students (Hypothesis 4) in determining intervention outcomes.

5.2. Participants

Sampling was purposive and multi-stage in nature. The Department For Education selected 10 Local Authorities that were representative of the diversity inherent in districts across England (in terms of population density, socio-economic factors, and geographical location). Each Local Authority purposively selected participant schools that represented the diversity of schools in the area in terms of factors such as attainment and ethnicity, albeit with a slight bias toward those with larger proportions of students with disabilities (to maximize the reach of the AfA pilot). Within each school, all students with
disabilities in Years 1 (aged 5–6), 5 (aged 9–10), 7 (aged 11–12) and 10 (aged 14–15) were selected to participate. These year groups were chosen by the Department for Education to provide broad representation of the different Key Stages of compulsory education in England.

5.2.1. Schools

Of the 431 mainstream primary and secondary schools that consented to participation in the study, 426 are represented in our analysis, made up of 351 primary and 75 secondary schools. This figure is lower than the overall number of participating schools as data were missing for a small number, and since our focus was on the attainment of students with disabilities in mainstream schools, data from participating special schools and referral units were excluded from the analysis. Table 1 presents a comparison of our school sample and all state-funded mainstream schools in England by size, attendance, attainment, free school meal eligibility, proportion of students speaking English as an additional language, and proportion of students with disabilities. National average comparison data was taken from statistical releases (Department for Education, 2010). Overall, schools did not differ greatly from national averages in terms of these contextual and compositional features, with the exception of secondary school size, where participating schools were somewhat larger than typical.

5.2.2. Students

Parental opt-out consent and student assent was sought for all participants. Baseline academic data was available for a total of 12,038 students with disabilities. However, complete datasets (e.g., baseline/pre-test and post-test) for English and/or Maths were available for N = 11,392, spread across Years 1 (n = 2291), 5 (n = 3048), 7 (n = 3086) and 10 (n = 2966) attending the above schools. Student characteristics for this sample can be seen in Table 2 and are compared to the national averages (Department for Education, 2010) to establish the representativeness of the sample.

5.3. Measures

5.3.1. Academic progress

Assessment of students’ academic progress in England utilizes distinct but overlapping metrics for children of different ages and abilities. First, Performance Scale Levels are used to assess children who are working below the first level of the National Curriculum (Qualifications & Curriculum Authority, 2009). Second, National Curriculum levels are used throughout Key Stages 1 to 3 (Qualifications & Curriculum Authority, 2010). Finally, General Certificate of Secondary Education gradings are used throughout Key Stage 4 (Department for Education, 2012). These metrics can be mapped converted to a generic “point score” using a system developed for the National Strategies Progression Guidance. This provides a single, continuous 3 Although the school-wide elements were, of course, expected to reach all students in a given school. 4 A Key Stage is a formal phase of state education in England. There are four Key Stages during the compulsory schooling period (1 = age 5–7; 2 = 7–11; 3 = 11–14; 4 = 14–16).
scale that spans all phases of compulsory education, from the beginning of primary school to the end of secondary school. The point score scale ranges from 1 (equivalent to Performance Scale Level 1) to 65 (equivalent to National Curriculum Level 10a or the General Certificate of Secondary Education grade A*+), with 4 points of progress being equivalent to 2 sub-levels of progress on the National Curriculum (which is the expected annual rate of progression for all students) (National Strategies, 2009). Teacher assessments in English and Maths were reported at baseline/pre-test (autumn 2009) and post-test (summer 2011), covering a total of 19 months (16 months for Year 1 students).

Academic data from end of Key Stage assessments in the National Pupil Database held by the Department for Education were used to calculate the average rates of academic progress for students in primary and secondary schools across England. We calculated the average amount of progress made by students nationally in primary schools by computing the difference in the point scores achieved in their end of Key Stage 1 and end of Key Stage 2 assessments (data available for c. 650,000 students without disabilities, and c. 80,000 with disabilities). For secondary schools, we applied the same procedure but with end of Key Stage 2 and end of Key Stage 4 assessment data (data available for c. 910,000 students without disabilities, and c. 100,000 with disabilities). Having calculated the average total progress made between Key Stages among students nationally, we worked out their average monthly progress and used this to estimate an average rate of progress over the 19-month period of the AfA pilot.

5.3.2. School and student characteristics

‘Background’ socio-demographic and other descriptive data at school (e.g. size) and individual (e.g. gender) levels were collected by the authors from various sources including the National Pupil Database (for individual level data), Edubase performance tables (www.edubase.gov.uk) and/or information provided by Local Authorities (for school level data).

5.3.3. Other data

AfA implementation data was collected using a school-level survey completed by the program coordinator in each participating school. Data pertaining to students’ behavior, bullying and positive relationships was collected using the Wider Outcomes Survey for Teachers (Wigelsworth, Oldfield, & Humphrey, 2013). Data was matched at individual and school levels using unique identification numbers that are used in student and school censuses in England.

Table 3 outlines each study variable, including level, scoring range and interpretation, and justification for inclusion.

---

5 Pre-test for Year 1 students was delayed until spring 2010 as many were still working at Foundation Stage levels – which cannot be converted reliably to the point score scale – in autumn 2009.

6 An obvious limitation of this method is that it assumes that academic progress is linear across a Key Stage(s), which may not be the case for all students.
<table>
<thead>
<tr>
<th>Level</th>
<th>Variable</th>
<th>Scoring protocol</th>
<th>Justification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local Authority</td>
<td>–</td>
<td>–</td>
<td>Ruling out any variation in outcomes attributable to differences between Local Authorities enhances power, precision and rigor of analysis</td>
</tr>
<tr>
<td>School</td>
<td>Compositional and contextual variables</td>
<td>% English as an additional language Proportion of students at school speaking English as an additional language, expressed as a percentage from 0 to 100</td>
<td>Aggregated compositional characteristics offer a proxy index of the diversity inherent in the student population. Such factors have been demonstrated to co-vary with attainment (Humphrey et al., 2013a,b) and may be important influences on intervention outcomes (for example, a higher proportion of students with disabilities in a given school may influence the responsive capacity of staff). Aggregated school outcome variables such as attendance and attainment provide broad indicators of school performance, which may influence intervention outcomes via program differentiation mechanisms (for example, the intervention may be seen as more distinct in a lower performing school). Indicators of school climate have been demonstrated to influence attainment (Battistich, Solomon, Kim, Watson, &amp; Schaps, 1995; Blau, Lamb, Stearns, &amp; Pellerin, 2001). For example, positive relationships and levels of bullying are both correlates of academic progress (Konishi, Hymel, Zumbo, &amp; Li, 2010)</td>
</tr>
<tr>
<td>School</td>
<td>% Free school meals</td>
<td>% Free school meals Proportion of students at school eligible for free school meals, expressed as a percentage from 0 to 100</td>
<td></td>
</tr>
<tr>
<td>School</td>
<td>% School Action Plus or with Statements of Special Educational Needs</td>
<td>% School Action Plus or with Statements of Special Educational Needs, expressed as a percentage from 0 to 100</td>
<td></td>
</tr>
<tr>
<td>School</td>
<td>% Achievement</td>
<td>% Achievement Proportion of students achieving at least National Curriculum Level 4 in English and Maths (primary) or at least 5 A*–C General Certificate of Secondary Education grades including English and Maths (secondary), expressed as a percentage from 0 to 100</td>
<td></td>
</tr>
<tr>
<td>School</td>
<td>% Absence</td>
<td>% Absence Average proportion of student absence from school (expressed as %, from 0 to 100)</td>
<td></td>
</tr>
<tr>
<td>School behavior problems</td>
<td></td>
<td>School behavior problems Aggregated student behavior scores at pre-test derived from the Wider Outcomes Survey for Teachers (0–3, higher scores indicating more behavior problems)</td>
<td></td>
</tr>
<tr>
<td>School positive relationships</td>
<td></td>
<td>School positive relationships Aggregated student positive relationships scores at pre-test derived from the Wider Outcomes Survey for Teachers (0–3, higher scores indicating better relationships)</td>
<td></td>
</tr>
<tr>
<td>School bullying</td>
<td></td>
<td>School bullying Aggregated student bullying scores at pre-test derived from the Wider Outcomes Survey for Teachers (0–3, higher scores indicating more bullying)</td>
<td></td>
</tr>
<tr>
<td>Implementation variables</td>
<td>Leadership</td>
<td>Leadership “Is AfA school lead a member of the School Leadership Team?” No (0) vs. Yes (1)</td>
<td>Leadership involvement is an important catalyst in school-based interventions (Leithwood, Harris, &amp; Hopkins, 2008) and has been shown to influence outcomes in other studies (e.g. Kam, Greenberg &amp; Walls, 2003) Effective implementation of Assessing Pupil Progress target-setting (high moderation, sharing with parents, regular review) is proposed as a mechanism of change in AfA (see Figure 1)</td>
</tr>
<tr>
<td></td>
<td>Assessing Pupil Progress moderation</td>
<td>Assessing Pupil Progress moderation “How are judgments surrounding assessment of student progress moderated?” Scale scored from 1 to 5, with higher scores indicating increased use of moderation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Assessing Pupil Progress sharing</td>
<td>Assessing Pupil Progress sharing “Are targets set using the progression guidance shared with parents?” No (0) vs. Yes (1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Assessing Pupil Progress target review teachers</td>
<td>Assessing Pupil Progress target review teachers “Are these targets reviewed by teachers on a termly basis?” Never/ Rarely/Often/Always</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Assessing Pupil Progress target review parents</td>
<td>Assessing Pupil Progress target review parents As above but for parents</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intervention planning</td>
<td>Intervention planning “Who is involved in planning interventions?” Scale scored from 1 to 5, with higher scores indicating larger numbers of people</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intervention form</td>
<td>Intervention form “What form do these interventions take?” Scale scored from 1 to 4, with higher scores indicating more comprehensive forms of intervention</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Information access</td>
<td>Information access “Who has access to student information?” Scale from 1 to 6, with higher scores indicating larger numbers of people</td>
<td></td>
</tr>
<tr>
<td>Parent communication</td>
<td>“How do parents get to know how their children are doing?” Scale scored from 1 to 8, with higher scores indicating more methods of communication</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parent relationships</td>
<td>“How would you describe the school’s relationship with parents prior to AfA?” Very Poor/Poor/Good/Excellent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structured conversation dosage</td>
<td>% of students at school from whom at least 2 structured conversations took place</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structured conversation fidelity</td>
<td>Extent to which schools reported following 9 key principles of structured conversations practice. Scale scored from 0 to 36, with higher scores indicating greater adherence to guidance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive relationships strategies implementation</td>
<td>Extent to which school reported progress in implementation of strategies to facilitate positive relationships. Scale scored from 0 to 3, with higher scores indicating greater implementation progress</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Behavior problems strategies implementation</td>
<td>As above but for behavior problems</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bullying strategies implementation</td>
<td>As above but for bullying</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attendance strategies implementation</td>
<td>As above but for attendance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wider participation strategies implementation</td>
<td>As above but for wider participation</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Student | Gender | Male/Female |
| Year group | Year 1/Year 5/Year 7/Year 10 |
| Ethnicity | White British, Asian, Black, Mixed, Chinese, Other, Unclassified |
| Language | English, Other, Unclassified |
| FSM eligibility | Not eligible (0) vs. eligible (1) |
| Special educational provision | School Action, School Action Plus, Statement of Special Educational Needs, Unclassified |
| Attendance | Proportion of days attending school, expressed as a percentage from 0 to 100 |
| Behavior | Student behavior scores at pre-test derived from the Wider Outcomes Survey for Teachers (0–3, higher scores indicating more behavior problems) |
| Positive relationships | Student positive relationships scores at pre-test derived from the Wider Outcomes Survey for Teachers (0–3, higher scores indicating better relationships) |
| Bullying | Student bullying scores at pre-test derived from the Wider Outcomes Survey for Teachers (0–3, higher scores indicating more bullying) |
| Baseline academic attainment | Maths and English point score at pre-test (range 1–65, with higher scores indicating greater achievement) |
| Response variable | Maths and English point score at post-test (range 1–65, with higher scores indicating greater achievement) |

Foundations of positive home-school collaboration (e.g., communication, relationships) are fundamental to the success of interventions involving parents (Lendrum, Barlow & Humphrey, 2013) and students’ academic progress (Fan & Chen, 2001).

High quality implementation of the Structured Conversations with Parents component of AfA was theorized to be a driver of change and thus improved outcomes for students (see Figure 1).

Progress in implementation of provision to promote wider outcomes may influence attainment. For example, interventions to promote social, emotional and behavioral outcomes have been demonstrated to impact upon academic progress (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011).

Student socio-demographic characteristics represent indices of individual differences that may predict variability in intervention outcomes (for example, student socio-economic background – Holsen, Iversen & Smith, 2009).

Disability type and special educational provision both co-vary with academic attainment (Humphrey et al., 2013a,b) and are likely to influence the intervention approach taken with targeted students.

Intervention exposure (e.g. reach, responsiveness) and therefore outcomes are likely to be significantly enhanced for students who attend school more frequently.

Theory and research in developmental cascades support the notion that student psychosocial outcomes influence later attainment (Masten & Cicchetti, 2010).

Inclusion of baseline/pre-test significantly increases power, precision and rigor of analysis.
5.4. Procedure

Key Teachers completed the Wider Outcomes Survey for Teachers online at pre-test. Implementation surveys were completed by schools’ AfA project coordinators between pre-test and post-test (April/May 2011). Assessment of students’ academic attainment in English and Maths was provided by the relevant subject teachers at pre- and post-test as noted above. This academic data was collated first at school level, where it was checked and verified by the school’s special education coordinator or equivalent, before being securely transferred to the relevant Local Authority, and then on to the National Strategies (representing the Department For Education), who performed data cleaning duties before handing the final dataset over to the research team via secure channels.

5.5. Analytical strategy

Two approaches to analysis were undertaken. First, for Hypotheses 1a, 1b, 2a, and 2b, we utilized one-sample t-tests (correcting for multiple comparisons) to assess the academic progress made by students with disabilities in AfA schools relative to those with and without disabilities in our national dataset. We examined the associated effect sizes in the context of the three benchmarks for practical significance outlined earlier. Second, for Hypotheses 3a, 3b and 4, we fitted fixed-effects, ‘intercept as outcomes’ (Raudenbush & Bryk, 2002) hierarchical linear models. Following Peugh (2010), we report both global (proportional reduction in variance in the response variable associated with the inclusion of all explanatory variables) and local (the influence of individual variables on the response variable) effects for these models.

6. Results

6.1. The impact of AfA on the academic progress in English of students with disabilities (H1a, H2a)

English data was available at pre-test for 12,038 students and at post-test for 10,996 of these (missing data for 8.7% of the sample). Missing tabulation pattern analysis showed the data was not Missing Completely at Random because a higher proportion of data on students’ special education provision and disability category were missing compared to other explanatory variables. However, as there was no indication that any of the values missing in the analysis were due to either of these variables, the data was considered Missing at Random (Heitjan & Basu, 1996). As a sufficient sample size was retained for inferential analysis, data imputation was not considered appropriate (McKnight, McKnight, Sidani, & Figueredo, 2007). Therefore, the analysis below is based on the complete datasets for N = 10,996 students with disabilities.

Mean progress in English during the AfA intervention period was 5.74 on the point score scale. Tables 4 and 5 provide details of the progress made by students in each year group, alongside comparative data for students with and without disabilities nationally over an equivalent period of time, and associated effect sizes. Given that the number of comparisons involved would increase family-wise error rate, Alpha was conservatively set at 0.006 using the standard Bonferroni formula (e.g. standard of 0.05 divided by 8 comparisons = 0.006). One sample t-tests demonstrated that students with disabilities attending AfA schools made significantly greater progress than their peers with and without disabilities nationally, the only exception being Year 7 students (who made significantly more progress than students with disabilities nationally, but significantly less progress than those without disabilities nationally).

These analyses speak to the success of the AfA program in significantly improving academic attainment in English of students with disabilities. Benchmark 1 (normative change) was exceeded in all cases. Benchmark 2 (policy-relevant gaps) was exceeded in all but one case (students in Year 7). Benchmark 3 (similar interventions) was exceeded in 2 of 4 cases. Using Cohen’s U3 index, all but one (Year 7) of our comparisons with progress of students with disabilities nationally produced effect sizes that met or exceeded Durlak’s (2009) criteria of 10% (d = 0.25) improvement as being of ‘practical significance’. Collectively, the analyses provide support for Hypotheses 1a and 2a (albeit with a slight exceptions, as noted above).

Table 4
Mean point score progress (SD) in English made by students in AfA schools and comparison students with disabilities nationally.

<table>
<thead>
<tr>
<th></th>
<th>AfA students with disabilities</th>
<th>Students with disabilities nationally</th>
<th>Cohen’s d</th>
<th>Swanson and Hoskyn’s (1998) mean effect size (0.45) exceeded?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 1</td>
<td>4.43 (2.56)*</td>
<td>4.11</td>
<td>0.25</td>
<td>×</td>
</tr>
<tr>
<td>Year 5</td>
<td>6.85 (2.84)*</td>
<td>4.88</td>
<td>1.39</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Secondary</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 7</td>
<td>5.01 (3.56)*</td>
<td>4.70</td>
<td>0.17</td>
<td>×</td>
</tr>
<tr>
<td>Year 10</td>
<td>6.66 (7.77)</td>
<td>4.70</td>
<td>0.50</td>
<td>✓</td>
</tr>
</tbody>
</table>

* p < 0.006 (Bonferroni correction) in one sample t-tests using national dataset as comparator.

NB: Point score scale ranges from 1 to 65.
Table 5
Mean point score progress (SD) in English made by students in AFA schools and comparison students without disabilities nationally.

<table>
<thead>
<tr>
<th></th>
<th>AFA students with disabilities</th>
<th>Students without disabilities nationally</th>
<th>Cohen's $d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 1</td>
<td>4.43 (2.56)$^*$</td>
<td>4.20</td>
<td>0.18</td>
</tr>
<tr>
<td>Year 5</td>
<td>6.85 (2.84)$^*$</td>
<td>4.99</td>
<td>1.31</td>
</tr>
<tr>
<td>Secondary</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 7</td>
<td>5.01 (3.56)$^*$</td>
<td>5.94</td>
<td>−0.52</td>
</tr>
<tr>
<td>Year 10</td>
<td>6.66 (7.77)$^*$</td>
<td>5.94</td>
<td>0.19</td>
</tr>
</tbody>
</table>

* $p < 0.006$ (Bonferroni correction) in one sample t-tests using national dataset as comparator.

NB: Point score scale ranges from 1 to 65.

6.2. The impact of AFA on the academic progress in Maths of students with disabilities (H1b, H2b)

Maths data was available at pre-test for 11,974 students and for 11,096 of these at post-test (missing data for 7.2% of the sample). Missing tabulation pattern analysis produced identical results to that reported above for the English dataset. The analysis below is therefore based on the complete datasets for $N=11,096$ students with disabilities.

Analysis of the impact of AFA on students’ progress in Maths followed the same procedures as outlined above. Mean progress in Maths during the AFA intervention period was 5.57 on the point score scale. Tables 6 and 7 provide details of the progress made by students in each year group, alongside comparative data for students with and without disabilities nationally over an equivalent period of time, and associated effect sizes. One sample t-tests demonstrated that students with disabilities attending AFA schools made significantly greater progress than their peers with and without disabilities nationally, the only exception being Year 7 students (who made significantly more progress than students with disabilities nationally, but significantly less progress than those without disabilities nationally).

These analyses demonstrate that the AFA program was successful in significantly improving academic attainment in Maths for students with disabilities. Benchmark 1 (normative change) was exceeded in all cases. Benchmark 2 (policy-relevant gaps) was exceeded in all but one case (students in Year 7). In 3 of 4 cases Benchmark 3 (similar interventions) was exceeded. Using Cohen’s U3 index, all comparisons with progress of students with disabilities nationally produced effect sizes that met or exceeded Durlak’s (2009) criteria of 10% ($d = 0.25$) improvement as being of ‘practical significance’. Collectively, the analyses provide support for Hypotheses 1a and 2a (albeit with marginal exceptions, as noted above).

Table 6
Mean point score progress (SD) in Maths made by students in AFA schools and comparison students with disabilities nationally.

<table>
<thead>
<tr>
<th></th>
<th>AFA students with disabilities</th>
<th>Students with disabilities nationally</th>
<th>Cohen’s $d$</th>
<th>Swanson and Hoskyn’s (1998) mean effect size (0.45) exceeded?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 1</td>
<td>5.03 (2.82)$^*$</td>
<td>3.81</td>
<td>0.87</td>
<td>✓</td>
</tr>
<tr>
<td>Year 5</td>
<td>6.49 (3.23)$^*$</td>
<td>4.52</td>
<td>1.22</td>
<td>✓</td>
</tr>
<tr>
<td>Secondary</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 7</td>
<td>4.73 (3.56)$^*$</td>
<td>4.07</td>
<td>0.31</td>
<td>×</td>
</tr>
<tr>
<td>Year 10</td>
<td>6.01 (8.11)$^*$</td>
<td>4.07</td>
<td>0.48</td>
<td>✓</td>
</tr>
</tbody>
</table>

* $p < 0.006$ (Bonferroni correction) in one sample t-tests using national dataset as comparator.

NB: Point score scale ranges from 1 to 65.

Table 7
Mean point score progress (SD) in Maths made by students in AFA schools and comparison students without disabilities nationally.

<table>
<thead>
<tr>
<th></th>
<th>AFA students with disabilities</th>
<th>Students without disabilities nationally</th>
<th>Cohen’s $d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 1</td>
<td>5.03 (2.82)$^*$</td>
<td>4.02</td>
<td>0.72</td>
</tr>
<tr>
<td>Year 5</td>
<td>6.49 (3.23)$^*$</td>
<td>4.77</td>
<td>1.07</td>
</tr>
<tr>
<td>Secondary</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 7</td>
<td>4.73 (3.56)$^*$</td>
<td>5.55</td>
<td>−0.39</td>
</tr>
<tr>
<td>Year 10</td>
<td>6.01 (8.11)$^*$</td>
<td>5.55</td>
<td>0.11</td>
</tr>
</tbody>
</table>

* $p < 0.006$ (Bonferroni correction) in one sample t-tests using national dataset as comparator.

NB: Point score scale ranges from 1 to 65.
6.3. Effects of variability in the implementation environment (H3a), implementation processes and practices (H3b), and student characteristics (H4) on intervention outcomes

6.3.1. English

Two fixed-effects, ‘intercept as outcomes’ (Raudenbush & Bryk, 2002) hierarchical linear models were produced (separate multi-level analyses for primary and secondary schools; see Tables 8 and 9). In each case, the first stage of the analysis involved the production of an unconditional model that enabled approximations of the proportional variance in the response variable attributable to each level. A conditional model was then executed, in which the various explanatory variables at the school and individual levels (see Table 3) were included. The fit to the data of the unconditional and conditional models was then assessed by computing and comparing their $2\log$likelihood values at convergence. This allowed us to infer that the inclusion of the explanatory variables in the full model improved predictive power to a statistically significant degree. In the interests of brevity, only those explanatory variables where $p < .10$ are reported in the tables.

The unconditional models suggested that differences between schools (10.1–19.2%) and individuals (80.4–86.7%) were both important in determining academic progress in English. Differences between Local Authorities were not significant. In line with Hypotheses 3a, 3b and 4, the addition of the school and student level explanatory variables improved the primary school model fit to a statistically significant degree, $\chi^2 (62, n = 1741) = 9788.72, p < .001$. The addition of the contextual/compositional and implementation variables explained 90.6% of school level variance, while the addition of socio-demographic and psychosocial variables explained 84.6% of the student level variance. The explanatory variables that contributed statistically significant amounts of variance at school and student levels are presented in Table 8. Consistent with reporting conventions for hierarchical linear modeling, each coefficient is unstandardized and represents the amount of change in the response variable (in this case, English point score) associated with changes of 1 unit in the given explanatory variable. At the student level, for example, students with Specific Learning Difficulties experienced a 0.4 point score increase when compared to the reference group of those with Moderate Learning Difficulties. To aid interpretation of this data, the reader is reminded that the English point score scale ranges from 1 to 65, and that the average change score in AfA primary schools was 5.82 (SD 2.97).

The findings for secondary schools were also consistent with Hypotheses 3a, 3b and 4. The addition of the school and individual level explanatory variables significantly improved model fit, $\chi^2 (62, n = 1239) = 9412.80, p < .001$. The school contextual/compositional and implementation variables explained all of the school level variance, while the addition of the

<table>
<thead>
<tr>
<th>$\beta_{00k}$ = 18.08 (0.275)</th>
<th>$\beta_{00k} = -0.327 (2.576)$ model</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Coefficient</strong> $\beta$</td>
<td><strong>Coefficient $\beta$</strong></td>
</tr>
<tr>
<td>Local Authority</td>
<td>0.000</td>
</tr>
<tr>
<td>School</td>
<td>0.851</td>
</tr>
<tr>
<td>% students eligible for free school meals</td>
<td>+0.019</td>
</tr>
<tr>
<td>% students achieving Level 4 in English/Maths</td>
<td>+0.033</td>
</tr>
<tr>
<td>Student</td>
<td>5.804</td>
</tr>
<tr>
<td>Year group</td>
<td>+0.019 (if ‘Year 5’)</td>
</tr>
<tr>
<td>Ethnic group (compared to white)</td>
<td>+0.522 (if ‘black’)</td>
</tr>
<tr>
<td></td>
<td>-2.466 (if ‘Chinese’)</td>
</tr>
<tr>
<td>Language group (compared to English)</td>
<td>+0.821 (if ‘other’)</td>
</tr>
<tr>
<td>Eligibility for free school meals (compared to ‘no’)</td>
<td>-0.412 (if ‘yes’)</td>
</tr>
<tr>
<td>Special educational provision (compared to ‘School Action’)</td>
<td>-0.384 (if ‘School Action Plus’)</td>
</tr>
<tr>
<td>Disability category (compared to ‘moderate learning difficulties’)</td>
<td>-1.957 (if ‘Statement’)</td>
</tr>
<tr>
<td></td>
<td>+0.402 (if ‘specific learning difficulties’)</td>
</tr>
<tr>
<td></td>
<td>-2.241 (if ‘severe learning difficulties’)</td>
</tr>
<tr>
<td></td>
<td>+1.194 (if ‘behavioral, emotional and social difficulties’)</td>
</tr>
<tr>
<td></td>
<td>+2.457 (if ‘unknown’)</td>
</tr>
<tr>
<td>Positive relationships score at baseline</td>
<td>+0.287</td>
</tr>
<tr>
<td>% attendance (09/10)</td>
<td>+0.018</td>
</tr>
<tr>
<td>English baseline score (1–65)</td>
<td>+0.882</td>
</tr>
</tbody>
</table>

$-2\log$likelihood = 17,917.456 $\chi^2 (62, n = 1741) = 9788.72, p < .001$
socio-demographic and psychosocial variables explained 60.9% of the student level variance. The explanatory variables that contributed statistically significant amounts of variance at school and student levels are presented in Table 9. The average point score change in AfA secondary schools was 5.82 (SD 6.07).

### 6.3.2. Maths

The analytical procedures described above were repeated for the Maths response variable (see models in Tables 10 and 11).

The unconditional models suggested that differences between both schools (3.8–17.8%) and individuals (82.1–95%) were important in determining academic progress in Maths. Differences between Local Authorities were not significant. In line with Hypotheses 3a, 3b and 4, the addition of the school and student level explanatory variables improved the primary school model fit to a statistically significant degree, $\chi^2 (62, n = 1764) = 9536.25, p < .001$. The addition of the contextual/compositional and implementation variables explained 92.5% of school level variance, while the addition of socio-demographic and psychosocial variables explained 80.4% of the student level variance. As with English, the Maths point score scale ranges from 1 to 65. The average change in AfA primary schools was 5.86 (SD 3.15).

The findings for secondary schools also provided support for Hypotheses 3a, 3b and 4. The addition of the school and individual level explanatory variables significantly improved model fit, $\chi^2 (62, n = 1207) = 9824.49, p < .001$. The school contextual/compositional and implementation variables explained all of the school level variance, while the addition of the socio-demographic and psychosocial variables explained 59.5% of the student level variance. The explanatory variables that contributed statistically significant amounts of variance at school and student levels are presented in Table 11. The average change in AfA secondary schools was 5.37 (SD 6.50).

Taken as a whole, the analyses of progress in both English and Maths in primary and secondary schools involved in the AfA program provided strong support for Hypotheses 3a, 3b and 4. At the global level, the proportional reduction in variance associated with the introduction of the range of explanatory variables was substantial and statistically significant. At the local level, the explanatory variables that were found to influence student outcomes were consistent with the AfA logic model and the broader body of theory and research outlined in Section 2.
Table 10  
Influences on progress in Maths during the AfA pilot (primary schools).  

<table>
<thead>
<tr>
<th></th>
<th>Coefficient $\beta$</th>
<th>SE</th>
<th>p</th>
<th></th>
<th>Coefficient $\beta$</th>
<th>SE</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local Authority</td>
<td>0.040</td>
<td>0.260</td>
<td>.442</td>
<td>School</td>
<td>0.000</td>
<td>0.000</td>
<td>.000</td>
</tr>
<tr>
<td>School</td>
<td>8.219 (0.074)</td>
<td>1.159</td>
<td>&lt;.001</td>
<td>% students eligible for free school meals</td>
<td>+0.024</td>
<td>0.010</td>
<td>.009</td>
</tr>
<tr>
<td></td>
<td>17.8%</td>
<td></td>
<td></td>
<td>% students achieving Level 4 in English/Maths</td>
<td>+0.037</td>
<td>0.009</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>% absence</td>
<td>+0.71</td>
<td>0.022</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Student</td>
<td>37.959 (0.013)</td>
<td>1.063</td>
<td>&lt;.001</td>
<td>Student</td>
<td>7.452 (92.4%)</td>
<td>0.261</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>82.1%</td>
<td></td>
<td></td>
<td>Year group</td>
<td>–0.842 (if 'Year 5')</td>
<td>0.068</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sex (compared to 'male')</td>
<td>–0.354</td>
<td>0.146</td>
<td>.008</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ethnic group (compared to 'White British')</td>
<td>+0.582</td>
<td>0.347</td>
<td>.047</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Language group (compared to 'English')</td>
<td>+0.612</td>
<td>0.311</td>
<td>.025</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Special educational provision (compared to 'School Action')</td>
<td>–0.698</td>
<td>0.170</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Disability category (compared to 'moderate learning difficulties')</td>
<td>–1.895</td>
<td>0.319</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Positive relationships score at baseline</td>
<td>+1.253</td>
<td>0.234</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>% attendance (09/10)</td>
<td>+0.040</td>
<td>0.011</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Maths baseline score (1–65)</td>
<td>+0.822</td>
<td>0.024</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

$-2\log$ likelihood = 18,179.855  
$-2\log$ likelihood = 8643.604  
$\chi^2 (62, n = 1764) = 9536.25, p < .001$  

7. Discussion  
The current study produced a number of noteworthy findings that have implications for theory, research and practice in relation to interventions for students with disabilities, and the associated field of school-based prevention.

7.1. Impact of AfA on the academic progress of students with disabilities  
Students in the AfA cohort made considerable progress in both English and Maths. In all cases this progress was significantly greater than their peers with disabilities nationally (Benchmark 1), and for students in Years 1, 5 and 10 it was greater than for those without disabilities nationally (Benchmark 2) over an equivalent period of time. The effect sizes associated with these differences ranged from small to very large, and in 5 of 8 cases exceeded the average effect on general academic achievement reported in Swanson and Hoskyn (1998) meta-analysis (Benchmark 3). Furthermore, in terms of meaningful change in the classroom, all but one case (Year 7 English) in the comparisons with progress of students with disabilities nationally produced effect sizes that met or exceeded Durlak’s (2009) criteria of 10% (d = 0.25) improvement as being of ‘practical significance’. In this sense, AfA can be considered successful in achieving its aim of improving outcomes for students with disabilities, especially when considered alongside our other findings, which produced similarly positive results (e.g., Humphrey, Wigelsworth, et al., 2013).

This aspect of our study also highlights the utility of Hill et al.’s (2008) framework for assessing practical significance. The current study is the first in the field to explicitly utilize these authors’ empirical benchmarks for interpreting intervention effect sizes in relation to the scaling up of a pilot intervention. Our findings were sufficient to convince the Department For Education to bring AfA to scale nationally (Department For Education, 2011) at a time when other initiatives were being withdrawn and £670 million (roughly $1.1 billion) was cut from the national education budget. Despite this, it is important to note that intervention effect sizes and their accompanying benchmarks need to also be considered in relation to associated costs. Harris (2008) proposes a useful approach to modeling the cost-effectiveness of interventions that could be incorporated into the aforementioned framework, thereby aligning policy analysis and decision-making in education with that of other disciplines such as health.

In spite of the above, the findings for students with disabilities in Year 7 consistently bucked the positive trend. This may reflect a variety of issues. First, Year 7 is the beginning of secondary school in England. Transition from primary to secondary school is a potentially difficult and stressful period for all students, due to curriculum discontinuity, a shift in academic
expectations and relationships with teachers, increased social demands, and so on. Those with disabilities are arguably more vulnerable in this regard, as they may find the experience generally more daunting (Reay & Lucey, 2000), exacerbated by reduced parental involvement and the cessation of (and changes to) interventions that were in place in primary school (Yeboah, 2002). It is therefore crucial that accurate records are passed on to the secondary school to ensure appropriate measures are in place to ease the transition process. Second, the older students – those in Years 5 (primary schools) and 10 (secondary schools) – had been at their respective schools for a longer period of time, meaning that teaching staff were presumably more aware of their specific needs and had a clearer sense of intervention approaches that would be more suitable and effective. These students are also at the latter phase of primary and secondary schooling respectively, and in preparing for end of Key Stage examinations and assessments they may have experienced a greater focus on raising achievement levels. Finally, the method used for calculating our national comparison figures (see ‘Measures’) assumes that progress is linear; that is, students make roughly equal progress from one year to the next. However, this is not always the case, and the Year 7 findings may actually reflect a curvilinear trend.

7.2. Effects of implementation environment, processes and practices

The current study also sought to understand how both aspects of the school environment and intervention processes and practices influence intervention outcomes. In terms of the former, two school contextual features appeared to be important, recurring in two of our four analyses. First, greater academic progress was made by students with disabilities in schools with higher proportions of students eligible for free school meals (our proxy for socio-economic status). Second, worse academic
progress was made in schools with higher proportions of students with disabilities. These findings are consistent with theory from implementation science (e.g., Domitrovich et al., 2008) and the limited number of empirical studies that have directly assessed the influence of school characteristics (e.g., Bierman et al., 2010; Pas, Waasdorp, & Bradshaw, 2014). An immediate implication of this is that when planning and implementing interventions like AfA, consideration needs to be given to the ‘make-up’ of the student body in order that resources are utilized effectively and that level of need does not outstrip a school’s capacity to respond. The findings also highlight the importance of accounting for compositional differences between schools in intervention research.

Unexpectedly, and in contrast to much of the implementation literature (ibid; Kam, Greenberg, & Walls, 2003), school leadership involvement did not influence intervention outcomes in any analysis. Our analyses of the Structured Conversations with Parents (Lendrum et al., 2013) and Developing Provision for Wider Outcomes (Humphrey, Wigelsworth, et al., 2013) strands of AfA demonstrated some effects to varying degrees for other outcomes. The lack of an association between school leadership involvement and outcomes in the Assessment, Tracking and Intervention strand may therefore indicate that it is more important as a catalyst for activity in the implementation of intervention practices that are often seen as less central to schools’ remit than academic intervention (e.g., developing provision for wider outcomes).

Other aspects of the implementation environment were identified as significant influences on intervention outcomes. Perhaps unsurprisingly, students with disabilities in schools with stronger parental relations prior to the implementation of AfA experienced better academic outcomes. This undoubtedly reflects the crucial role played by parents in the intervention (Lendrum et al., 2013), while also aligning with the broader body of literature that establishes the importance of parental engagement and involvement more generally (Barlow & Humphrey, 2012). Rather unexpectedly, higher baseline positive relationships at school-level were negatively associated with academic outcomes at post-test in two of our analyses. We expected that this proxy indicator of the social-emotional climate of participating schools would serve as a foundation for greater success in the achievement of expected outcomes. One possible explanation is that the theorized synergistic effects of AfA are more pronounced in schools where relationships are problematic. Hence, schools where this aspect of the climate was already strong may have experienced less benefit.

Several implementation processes and practices influenced intervention outcomes. First, the conduct of the Structured Conversations with Parents strand of AfA (e.g., increased fidelity, higher exposure/dosage) influenced academic progress in two of our four analyses. This finding offers support for a key proposition in both the AfA logic model and Domitrovich et al.’s (2010) integrated prevention model – namely, that independent strategies that are fused into a coherent intervention framework in a coordinated manner will produce a synergistic effect. Further validation is offered in case studies of implementation reported in our main report to the Department For Education, where practitioners considered the three AfA strands to be deeply intertwined in the way in which they influenced outcomes for students with disabilities (Humphrey & Squires, 2011a,b).

Second, key aspects of the delivery of the AfA’s Assessment, Tracking and Intervention strand were shown to be critical. Specifically, increased frequency of teacher/parent involvement in reviewing students’ academic targets and a larger range of methods used to communicate information about progress were associated with significantly better student-level outcomes. These findings are indicative of the merits of an ‘adaptive’ approach to academic intervention for students with disabilities. That is, what may ultimately be important in determining outcomes is the efficiency with which schools review targets, communicate information about progress and adapt approaches to intervention accordingly (e.g., where academic data is collected with greater frequency, this allows schools to assess and discuss whether a given approach is producing its intended outcomes, and make changes where this is not the case). This is akin to the Response to Intervention approach utilized in the United States (Fuchs & Fuchs, 2006). Interestingly, these findings applied in secondary but not primary schools. We speculate that this reflects the more considerable variation found in the former rather than these practices being unimportant in the latter.

These findings and those reporting on the other strands of AfA (e.g., Humphrey, Wigelsworth, et al., 2013; Lendrum et al., 2013) have enabled us to draw together some tentative ‘active ingredients’ of the program, thereby beginning to address recent concerns raised in relation to implementation science (Blase & Fixsen, 2013). Each of the following aspects of implementation were identified as predicting outcome variability at least two outcomes/analyses in the current study and/or associated papers (e.g., Humphrey, Wigelsworth, et al., 2013):

1. AfA lead is a member of the school leadership team,
2. High fidelity and dosage in structured conversations with parents,
3. More frequent involvement of class/subject teachers in reviewing individual targets,
4. More frequent involvement of parents in reviewing individual targets,
5. More comprehensive range of academic interventions used,
6. Larger range of individuals who have access to student data,
7. Larger range of methods used to communicate information about progress.

Of note here is the fact that no critical components were identified from the Developing Provision for Wider Outcomes strand of the program. This (non)finding could be attributed to program theory (e.g., the logic model is unsound),
implementation (e.g., the logic model is sound but the implementation of it was not), and/or evaluation (e.g., logic model and implementation were sound, but we were unable to measure them with sufficient precision). On balance, we feel that the latter is the most likely explanation. The Developing Provision for Wider Outcomes strand is the most flexible aspect of AfA and so capturing variation in practice was all the more challenging, particularly in comparison to the more prescriptive elements (e.g., the Structured Conversations with Parents strand). This issue has been reported elsewhere recently in relation to other flexible intervention frameworks (e.g., Deighton et al., in press). Given the general shift away from manualized, programmatic approaches in the field, this is an emergent methodological challenge that needs to be addressed in the study of implementation.

7.3. Differential gains among student subgroups

Aligning with notions of differential susceptibility to environmental influence in human development (e.g., Belsky & Pluess, 2009), individual differences between students with disabilities were also important in explaining variability in intervention outcomes. As we predicted, responsiveness to intervention was not uniform across the sample. Groups of learners who experienced accelerated progress in at least two analyses included those who attended school more regularly, those with stronger positive relationships, and those with behavioral, emotional and social difficulties or autism spectrum disorders.

In terms of attendance, this may been seen as a logical proxy for intervention reach (e.g., students have to attend school regularly in order to receive sufficient exposure to a given intervention; those who fail to attend school regularly are therefore less likely to experience its benefits). The influence of students’ positive relationships is perhaps unsurprising, as learning is an inherently social process. The fact that this variable influenced academic progress after controlling for prior attainment provides robust support for its hypothesized fortifying effects. This kind of effect is predicted by developmental cascade theory, which proposes that functioning in different domains is linked across time (Masten & Cicchetti, 2010), and in particular the body of work that focuses on cascades of positive youth development (e.g., one good thing leads to another – Lewin-Bizan, Bowers, & Lerner, 2010). In terms of disability category, the accelerated progress of those with behavioral, emotional and social difficulties and autism spectrum disorders may reflect advances in teacher training and development that were introduced concurrently with AfA (for example, the national Inclusion Development Program included bespoke content on these disability categories – Lindsay et al., 2011).

Groups of learners who experienced relatively lower levels of academic progress in the AfA pilot included females, those at the School Action Plus or Statement of Special Educational Needs stages of provision, and those with severe learning difficulties. In relation to the latter, this is perhaps unsurprising given the acute difficulties in cognition and learning that are characteristic of students to whom this disability category is applied. It may simply be the case that the AfA intervention model described in the current study was not sufficiently intensive trigger meaningful change for them. The same might also be said for those in receipt of provision at School Action Plus or with Statements of Special Educational Needs. The relatively lower levels of progress made by females is, however, less easily explainable. One possibility is that the well-established gender imbalance in most disability categories (Department for Education, 2014) led to a skewed intervention focus among teachers that favoured boys, at the expense of girls. Taken together, these findings have immediate applied value, inasmuch as they highlight groups of students who may require additional, more intensive support that augments the provision available in AfA (Humphrey & Squires, 2011a,b).

7.4. Limitations

In the interests of transparency it is important to acknowledge two key limitations in the current study. First, although we were able to compare the academic progress of students with disabilities in the study cohort with their peers with and without disabilities nationally, we were unable to utilize a ‘traditional’ control group in this strand of our evaluation. Furthermore, allocation to the AfA program was not randomized. Ultimately, this perhaps reflects the realities of some government-sponsored research projects in education, where implementation decisions are made that place limits upon what research designs are feasible (see Slee et al., 2009, for a similar example). Furthermore, research conducted in the ‘swampy lowlands’ of real-life practice must always strike a balance between rigor and relevance (Marshall, 2013), and what was sacrificed here was arguably outweighed by benefits accrued through increased ecological validity.

Second, our measures of academic progress of students with disabilities were based upon un-blinded teacher assessments as opposed to standardized tests. This was a necessary aspect of our design – it was simply infeasible to administer standardized tests to more than 12,000 students. We are reassured by the fact that teachers in participating schools had received training in the principles and practice of Assessing Pupil Progress (Department for Children Schools and Families, 2008c), which has been found to increase, “teachers’ accountability and the accuracy and consistency of assessment practice” (OFSTED, 2011, p.2). As above, ecological validity remains an important consideration. In England, students are only subjected to standardized tests at the end of key stages of education. In the interim, assessment of progress is determined by teacher judgment.
8. Conclusion

The aim of the current study was to document the implementation and impact of a novel intervention for students with disabilities in England: Achievement for All. Above and beyond establishing the effectiveness of this approach, we sought to contribute to the body of knowledge on, "not only … 'what works', but also about why various programs do or do not work, for whom and under what conditions they work, what is needed to scale up proven programs, and what policy supports are needed to scale them up without losing their effectiveness" (Slavin, 2012, p.xv). The analyses reported herein have provided some answers to these critical questions, while also establishing the practical significance of an original framework for intervention that seeks to improve the experiences and outcomes of the most vulnerable students in our education system.

Acknowledgments

This research was funded by the Department for Education, England. The authors thank the schools, staff and pupils involved in Achievement for All for their cooperation with and input into the evaluation.

References


